

FINAL
11-43-62
OCIT
0723

Subspace Based Signal Analysis of Partially Polarized Light Reflected by Plant Canopies

Dr. Fred L. Fontaine

The Cooper Union for the Advancement of Science & Art, New York, NY

December 23, 1996

Abstract

Final Report for NASA-Ames University Consortium NCC2-5101.
Collaborator: Dr. Vern C. Vanderbilt.

1 Introduction

The problem of concern here is the analysis of partially polarized light reflected by plant canopies. The significance of such signals and certain underlying mathematical models are presented in [3],[4],[5], for example.

We start with data measurements obtained at specified grid points with instruments with a particular field of view (FOV). The goal is to generate equivalent data values which would correspond to a more desirable array of grid points, such as uniformly spaced ones, and possibly with a different FOV. Then, polarization parameters, specifically the polarized reflectance R_{QU} and angle of the polarization plane χ , are computed. These quantities are related to the quantities R_Q and R_U which are linearly related to raw sensor measurements as:

$$R_{QU}e^{j2\chi} = R_Q + jR_U \quad (1)$$

The basic approach is to model the measurement process as linear, based on underlying "detail", obtaining an estimate of the detail and recomputing "simulated" measurements on the new grid with the new FOV. Actually, these two stages are combined into one so the dense "detail" is not

actually computed. Because of imperfections in the measurement process, least-square methods are used. The criterion for matching is a quadratic one directly on the data values themselves, or, alternatively, a quadratic one based on a nonlinear transformation of the measured data.

The solution formula in the linear case involves pseudo-inverses of matrices, and hence is computed using singular-value decomposition (SVD) methods.

The primary physical processes underlying the measurement process are:

- the blurring effect of a finite FOV;
- the variability of the area of the ground “footprint” with view angle;
- leaf size and orientation, ground cover and heliotropy; in particular, the solar position is not constant for all data measurements, and heliotropism introduces a “bias” in the probabilistic model of the leaf orientations;
- the different sensors were not collocated, so measurements taken at the same time at the same view angle result in partially overlapping “footprints”;
- the data fusion effect; that is, whether the data from different sensors, at different times, or at different wavelengths should be treated separately or combined, for example averaged;
- the desired polarization information, such as R_{QU} , is computed as a two step process: the first is a linear transformation of information from three different sensors (yielding R_Q , R_U) and the second is a nonlinear one ($R_{QU} = \sqrt{R_Q^2 + R_U^2}$, $\chi = \tan^{-1}(R_U/R_Q)/2$). Whether the linearly dependent parameters are recovered, then the nonlinearity is applied, or if the nonlinearly dependent parameters are computed directly does effect the ultimate outcome.

Whatever the details in the recovery method, the results must then be compared against those predicted by theoretical models. These theoretical models ultimately depend on models of probability density functions for various parameters, such as leaf orientation and ground cover, and their relation to source and view angle.

In the linear approach, these issues are addressed in the following fashion.

- the measurement process is a linear one which is relatively localized; hence it can be represented as a matrix acting on a data vector.
- the FOV concept yields a sparse matrix; if a single matrix is applied to the entire canopy area, the structure is not simple (for example, not banded) because of the way data is organized in the vector, corresponding to irregular 2-D sampling; alternatively, for every view angle for which a computation is to be performed, a full, lower dimensional matrix can be used to model the local measurement process and pseudo-inversion methods applied locally; this latter approach can be taken as a spatially varying by linear filtering process which must be inverted;
- data fusion can be handled by either treating data sources separately, or combining them into a joint vector;
- the linear transformation step in computing R_Q, R_U can be absorbed in the measurement matrix; however, this requires dealing with the three sensors together, and hence can cause inaccuracies because the sensors are not collocated.

In these notes, a linear least-squares method is developed, followed by one based on a nonlinear cost function. The methods were implemented using MATLAB for Windows on a Pentium PC and applied to data collected from a Sunflower field in 1991 [2]. The approach here is an outgrowth of the one developed in [1]. Several graphical results, shown in the Appendix, are discussed.

2 Linear Least-Squares Method

2.1 General Notation

In what follows, let $\mathcal{N}(B)$ and $\mathcal{R}(B)$ denote the nullspace and range space of a matrix B , and $B^\#$ the pseudo-inverse of B . Every matrix B is a one-to-one onto linear map from $\mathcal{N}^\perp(B)$ to $\mathcal{R}(B)$, and $B^\#$ is the inverse map. In particular, $Bx \in \mathcal{R}(B)$ and $B^\#y \in \mathcal{N}^\perp(B)$ for all x, y . The matrix $B^\#B$ is the orthogonal projection onto $\mathcal{N}^\perp(B)$, and $I - B^\#B$ is the orthogonal projection onto $\mathcal{N}(B)$. Also, $BB^\#$ is the orthogonal projection onto $\mathcal{R}(B)$. Let $\|\cdot\|$ denote the L^2 -norm of a vector, and the spectral norm of a matrix.

Let B^* denote the adjoint of B : the transpose for a real matrix, the Hermitian (conjugate) transpose for a complex matrix.

2.2 The Measurement Process

We first model the measurement process. By default, all data vectors are column vectors and linear operators are represented by matrices acting on the left. Let c denote the measured data, obtained at specified grid points with instruments with a particular field of view (FOV). Thus, let x represent the “underlying detail,” and the smoothing and sampling operations by the matrix A , resulting in the measurement process:

$$c = Ax + n \quad (2)$$

where n is the measurement noise. In general, A is rectangular and has fewer rows than columns. Let \mathcal{C} be the space of all measurement vectors c , and \mathcal{X} the space of all detail vectors x .

Our goal is to compute a data vector, d , corresponding to simulated measurements onto another set (say, a uniform pattern) of grid points and possibly at another FOV. The simulated measurement process is characterized by the matrix B . If x were known then:

$$d = Bx \quad (3)$$

Let \mathcal{D} denote the space of all simulated measurements d .

2.3 The Criterion

We first set up a quadratic criterion which yields a linear solution. Given c , the minimum-norm least-square solution for x minimizes $\|c - Ax\|$; since A does not have full column rank, in general, there are infinitely many minimizers, so the specific solution is one which, among all minimizers, also has minimum $\|x\|$. Denoting this solution by x_o , we have:

$$x_o = A^\# c \quad (4)$$

and more generally all minimizers of $\|c - Ax\|$ are given by:

$$x = A^\# c + x_n \quad (5)$$

where $x_n \in \mathcal{N}(A)$:

$$Ax_n = 0 \quad (6)$$

Let Q be the orthogonal projection onto $\mathcal{N}(A)$:

$$Q = I - A^\# A \quad (7)$$

Now, $Q^\# = Q^* = Q$. Moreover, $x_n \in \mathcal{N}(A)$ iff it satisfies:

$$x_n = Qx_n \quad (8)$$

The choice of x_n is arbitrary given only the information in c , and hence its effect on computed values constitutes artifacts.

Simulated measurements d based on such x have the form:

$$d = Bx = BA^\# c + Bx_n = BA^\# c + BQx' \quad (9)$$

where x' is arbitrary; that is, given c , then any vector of the form:

$$d = BA^\# c + d_o \quad (10)$$

where $d_o \in \mathcal{R}(BQ)$, that is d_o which is a linear combination of the columns of BQ , corresponds to minimizing $\|c - Ax\|$. The choice of $x_n = 0$, that is $x = x_o = A^\# c$ achieves minimum $\|x\|$ but *not* minimum $\|d\|$.

Smoothness (regularity) considerations suggest we may attempt to find a solution which minimizes $\|d\|$, not the “underlying detail,” which is never directly computed. We can take this one step further. Assume a-priori information in the form of a “nominal” d_1 is available, so we want to minimize $\|d - d_1\|$.

Thus, the problem is as follows: given c and d_1 (which is 0 if there is no a-priori information), pick d corresponding to a vector x (via $d = Bx$) which minimizes $\|c - Ax\|$ and, among all such, $\|d - d_1\|$ should be minimized.

Let us denote $\mathcal{R}(BQ)$ as the *artifact space*. It consists of all vectors in the simulated measurement space which arise from measurements performed on “underlying detail” which is in the *nullspace* of the actual measurements. Given an actual measurement c , there is no (direct) information on this null component, and hence the corresponding component of the simulated measurement d can be considered an *artifact*.

The orthogonal projector onto the artifact space is P given by:

$$P = (BQ)(BQ)^\# = UU^* = \sum_{k=1}^r u_k u_k^* \quad (11)$$

where $U = \begin{bmatrix} u_1 & u_2 & \cdots & u_r \end{bmatrix}$ is the $L \times r$ matrix whose orthonormal columns, the right singular vectors of BQ , span the artifact space. Note that $r = \text{rank}(BQ)$ is the dimension of this artifact space.

Theorem 1 *Given measurements c and a-priori information d_1 , the simulated data d which corresponds to minimum $\|d - d_1\|$ subject to the constraint that $d = Bx$ for some x minimizing $\|c - Ax\|$ is given by:*

$$d = (I - P)BA^\#c + Pd_1 \quad (12)$$

That is, the artifact component is computed directly from d_1 , and the component orthogonal to the artifact component is computed directly from c .

Proof. We have $x = A^\#c + Qx_o$, with x_o arbitrary, the complete list of x which minimize $\|c - Ax\|$. Then:

$$d = Bx = BA^\#c + BQx_o \quad (13)$$

Then:

$$d - d_1 = (BA^\#c - d_1) + BQx_o \quad (14)$$

A choice of x_o which minimizes:

$$d_1 - BA^\#c - BQx_o \quad (15)$$

is:

$$x_o = (BQ)^\#(d_1 - BA^\#c) \quad (16)$$

and corresponding d is:

$$d = BA^\#c + Pd_1 - PBA^\#c \quad (17)$$

The vector d in (12) corresponds to detail x which is a least-squares match to the measurements c , while minimizing error from a-priori information d_1 . But in the case of no a-priori information, we zero out the artifact component via:

$$\begin{aligned} d &= (I - P)BA^\#c \\ &= \left(I - B(I - A^\#A)(B(I - A^\#A))^\# \right) BA^\#c \end{aligned} \quad (18)$$

This involves two pseudo-inverse and five matrix multiplication operations; matrix addition and matrix-vector operations are of significantly lower computational complexity.

Localizing this reduces complexity to solve for a particular point or region, but must be repeated as the matrices need to be reformulated because of the irregularity of the original measurements.

3 Nonlinear Cost Function

The underlying cost functions in the previous sections are L^2 -norms, specifically $\|d - d_1\|$. As an alternative, we may try to match results computed from d , rather than d itself. For example, suppose we try to match R_{QU} to some a-priori knowledge or probabilistic model. This is a nonlinear function of a vector of low-dimension, typically three or less (here, R_{QU} is a nonlinear function of R_Q and R_U).

Denote this as:

$$\rho(\vec{\xi}) \quad (19)$$

where:

$$\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_N) \quad (20)$$

where $1 \leq N \leq 3$, typically. Hence, d and c are $L \times N$ and $M \times N$ matrices, respectively, with each row (length N) consisting of one vector for which ρ can be computed. Let:

$$\begin{aligned} c &= \begin{bmatrix} c_1 & c_2 & \dots & c_M \end{bmatrix}^T \\ d &= \begin{bmatrix} d_1 & d_2 & \dots & d_L \end{bmatrix}^T \end{aligned} \quad (21)$$

where each c_i and d_i is an N -dimensional column vector. We set up the cost function:

$$J(d, d_1) = \sum_{i=1}^L [\rho(d_i) - \hat{\rho}_i]^2 \quad (22)$$

where $\hat{\rho}_i$ is ρ applied to the i^{th} row of d_1 . We then find d of the form:

$$d = BA^\#c + BQx_o \quad (23)$$

which minimizes (22). Now, we can solve this constrained minimization problem using the method of Lagrange multipliers. However, we take an alternate route here. Instead, we obtain an explicit expression for d satisfying the constraints in terms of a minimal set of arbitrary parameters, and substitute it into the formula (22) for the cost function J . The unconstrained minimization problem can then be solved directly.

We can rewrite the constraint as:

$$(I - P)(d - BA^\#c) = 0 \quad (24)$$

or, in other words, $d - BA^\#c$ must be in the range space of BQ , that is an artifact component. Let $\{s_i\}_{i=1}^r$ be orthonormal vectors which span the artifact space, so that $P = \sum_{k=1}^r u_k u_k^* = SS^*$, where $U = \begin{bmatrix} u_1 & u_2 & \cdots & u_r \end{bmatrix}$ is the $L \times r$ matrix whose orthonormal columns span the artifact space. Note that $r = \text{rank}(BQ)$ is the dimension of this artifact space. Hence:

$$d = d_o + \sum_{k=1}^r u_k \gamma_k \quad (25)$$

where:

$$d_o = BA^\#c \quad (26)$$

is $L \times N$ is the simulated data computed without regard for artifacts, and each γ_k is $1 \times N$. The $r \times N$ elements in the matrix Γ given by:

$$\Gamma = \begin{bmatrix} \gamma_1^T & \gamma_2^T & \cdots & \gamma_M^T \end{bmatrix}^T \quad (27)$$

are free parameters to be chosen so as to minimize J :

$$J(\Gamma) = \sum_{i=1}^L \left[\rho \left(d_{oi} + \sum_{k=1}^r u_{ik} \gamma_k \right) - \hat{\rho}_i \right]^2 \quad (28)$$

where u_{ik} is the i^{th} component of the vector u_k , d_{oi} is the i^{th} row of the matrix d_o and $\hat{\rho}_i$ is the nominal or a-priori value of $\rho(d)$ to be matched.

For $1 \leq j \leq N$, let

$$\mu_j(\vec{\xi}) = \partial \rho(\vec{u}) / \partial \xi_j \quad (29)$$

Then the partial derivative of $J(\Gamma)$ with respect to the m_j^{th} element γ_{mj} of Γ is:

$$\frac{\partial J}{\partial \gamma_{mj}} = 2 \sum_{i=1}^L \left[\left(\rho \left(d_{oi} + \sum_{k=1}^r u_{ik} \gamma_k \right) - \hat{\rho}_i \right) \cdot \mu_j \left(d_{oi} + \sum_{k=1}^r u_{ik} \gamma_k \right) \cdot u_{im} \right] \quad (30)$$

and the solution to the minimization problem is achieved by setting this expression to 0 for $1 \leq m \leq N$, $1 \leq j \leq r$.

We can simplify the notation at this point somewhat. We can define the $N \times L$ matrix G via:

$$g_{ji} = \left(\rho \left(d_{oi} + \sum_{k=1}^r u_{ik} \gamma_k \right) - \hat{\rho}_i \right) \cdot \mu_j \left(d_{oi} + \sum_{k=1}^r u_{ik} \gamma_k \right) \quad (31)$$

where $1 \leq j \leq N$, $1 \leq i \leq L$. Then (30) reduces to:

$$(\partial J / \partial \Gamma)^T = 2GU \quad (32)$$

where $(\partial J / \partial \Gamma)^T$ is the $N \times r$ matrix whose jm^{th} element is $\partial J / \partial \gamma_{mj}$. Hence, the minimization problem is:

$$GU = 0 \quad (33)$$

This is a system of $N \cdot r$ equations (since GU is an $N \times r$ matrix) in $N \cdot r$ unknowns, the γ_{mj} 's.

Some comments at this point is warranted. In general, L may be large, corresponding to the number of local or data points being considered. However, N is generally small, maybe 2, and r , the rank of the artifact space, is selectable. It can be determined either by numerical studies of P , or bounded by an arbitrary bound on the computational procedure to be followed. It is generally easier to minimize a single nonlinear function of several variables rather than solving simultaneous nonlinear equations. We can do this by converting the system (33) into a single function:

$$J'(\Gamma) = \sum_{m=1}^N \sum_{k=1}^r (GU)_{mk}^2 \quad (34)$$

which is a nonnegative definite form, which achieves its minimum, 0, exactly when (33) is satisfied. We can rewrite (34) as:

$$J'(\Gamma) = \text{trace}(GU(GU)^*) = \text{trace}(GUU^*G^*) \quad (35)$$

and, since $P = UU^*$, we get the final form:

$$J'(\Gamma) = \text{trace}(GPG^*) \quad (36)$$

We have J' as a nonnegative definite nonlinear quadratic form, and its minimization yields the desired solution.

The key step here is the choice of r , which is tantamount to a subspace selection.

Example 2 Let the underlying parameters be R_Q and R_U , and $R_{QU} = \sqrt{R_Q^2 + R_U^2}$. Then $N = 2$, $\rho(\xi_1, \xi_2) = \sqrt{\xi_1^2 + \xi_2^2}$, and:

$$\begin{aligned} \mu_1(\xi_1, \xi_2) &= \xi_1 / \rho(\xi_1, \xi_2) \\ \mu_2(\xi_1, \xi_2) &= \xi_2 / \rho(\xi_1, \xi_2) \end{aligned} \quad (37)$$

Example 3 *If instead we work with the angle of polarization $\chi = \tan^{-1}(R_U/R_Q)/2$. Then $\rho(\xi_1, \xi_2) = \tan^{-1}(\xi_2/\xi_1)/2$ and:*

$$\begin{aligned}\mu_1(\xi_1, \xi_2) &= -\xi_2/2 (\xi_1^2 + \xi_2^2) \\ \mu_2(\xi_1, \xi_2) &= \xi_1/2 (\xi_1^2 + \xi_2^2)\end{aligned}\tag{38}$$

4 Application to Experimental Data and Results

The linear least-squares method and the method based on the nonlinear cost function were both applied to data collected from a sunflower field in [[2]]. The experiment involved two types of instruments: Barnes, which had a FOV of 1° , and Cimel, which had a FOV of 1° or 12° , depending on the day upon which measurements were performed. Results for data collected on 07/25/91, shown in the Appendix, are representative; on this date, the Cimel FOV was 12° .

The measurements were taken from a $13.5m$ tower in the center of a sunflower field approximately $100m \times 100m$ in sunlight. The polarized component of the measured reflectance is modelled [3],[4],[5] as caused by sunlight reflected specularly from the leaf surfaces, and hence Fresnel equations for specular reflection apply. In particular, we are interested in measuring the total polarized reflectance R_{QU} and the angle of the plane of polarization, χ .

The measurements were taken over multiple zenith angles for several azimuths, approximately ranging from 0° to 360° in increments of 45° . The angles of measurement were very irregular. In order to apply the methods discussed here, the measurement process had to be modelled by certain operators, specifically A and B .

An operator is developed for a fixed ray with a specified view azimuth and a zenith angle ranging from 0 to 90° ; actually, the zenith angle was restricted to a subrange since very small and very large zenith angles caused spots on the ground (due to the finite FOV) which either were directly below the tower or beyond the edge of the field. Each raw measurement c , underlying x and simulated d is obtained from light collected within a finite FOV. The locations of the spots at which c is measured is determined by the experimental data itself, and is generally irregular; multiple measurements from the same spot also can occur. The locations of the underlying detail x were on a dense grid

within the vicinity of this strip, and of d on a regular range of spots for the fixed azimuth. The azimuth of the x spots varied to encompass the width of the spot measurements of c and d . The underlying detail x was assumed to arise from a FOV of 1° .

Given a position of x , it corresponds to a spot whose center, length and width depend on the azimuth and zenith values; this spot partially overlaps a corresponding spot of c or d , respectively. Hence, each row of A and B , respectively, implements a weighted sum of x values to obtain final c and d values, respectively, with the weight proportional to the area of the x spot which lies within the c (d) spot. Hence, the averaging effect of the FOV for data on one grid set is converted to a possibly different FOV on another grid set by weighting linearly according to the area overlap of the spots. If two points have a large area overlap, the information is expected to be more highly correlated than between two points with a smaller area overlap. Distance between points is accounted for indirectly, in that spots which are further apart have a smaller common area. However, the area overlap method is reasonable considering the physical situation, a multitude of plant leaves over a ground cover.

The results, shown in the Appendix, show that the raw measurements of R_{QU} and χ are irregular, both for Barnes instruments with FOV 1° and Cimel instruments with FOV 12° . The linear least-squares method provides generally good matching results, which are significantly smoothed. For the case of the Barnes instruments, the R_{QU} values are somewhat smaller. This may reflect averaging over complex values $R_Q + jR_U$, then taking magnitude; phase cancellation due to noise causes diminished R_{QU} . The nonlinear method yields R_{QU} results similar to the linear method. The key to the nonlinear method is the a-priori information $\hat{\rho}$; further studies should use the mathematical models developed in, for example, [3],[4],[5], and perhaps quantities other than $\rho = R_{QU}$ for better results. The nonlinear method is still promising precisely because it more easily accommodates such an approach.

The values of χ obtained for both Barnes and Cimel were very good, surprisingly good considering the high sensitivity of phase information in many situations. The nonlinear approach produced very small χ values, suggesting recovered R_U values are too small relative to R_Q . This may also be caused by the matching to R_{QU} in the nonlinear method; perhaps combining matching to R_{QU} and χ would yield better results.

In summary, the methods developed here are very promising. Further investigations and development of specific implementation in MATLAB, and

particularly a more extensive utilization of established mathematical models of the underlying physical process, for example through the nonlinear cost function approach, is indicated.

References

- [1] F.L. Fontaine and V.C. Vanderbilt, "Signal Analysis of Partially Polarized Light Reflected by Light Canopies," *1992 NASA-ASEE Summer Fellowship Final Report*, Ecosys. Sci. Tech. Branch, NASA Ames, CA, 1992.
- [2] V.C. Vanderbilt, "Tower Sunflower Polarization Experiment," *Experiment No. 212-916403*, Montfavet INRA, France, 1991.
- [3] V.C. Vanderbilt, K. deVenecia, "Specular, Diffuse and Polarized Imagery of an Oat Canopy," *IEEE Trans. Geosci. Remote Sensing*, vol. 26, no. 4, July 1988, pp.451-462.
- [4] V.C. Vanderbilt, L. Grant, S. Ustin and R. Myneni, J. Ross (eds.), "Polarization of Light by Vegetation," Ch. 7 in *Photon-Vegetation Interactions*, Springer-Verlag, 1991, pp. 191-228.
- [5] V.C. Vanderbilt, L. Grant, "Plant Canopy Specular Reflectance Model," *IEEE Trans. Geosci. Remote Sensing*, vol. 23, no. 5, Spet. 1985, pp.722-730.
- [6] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, 1995.

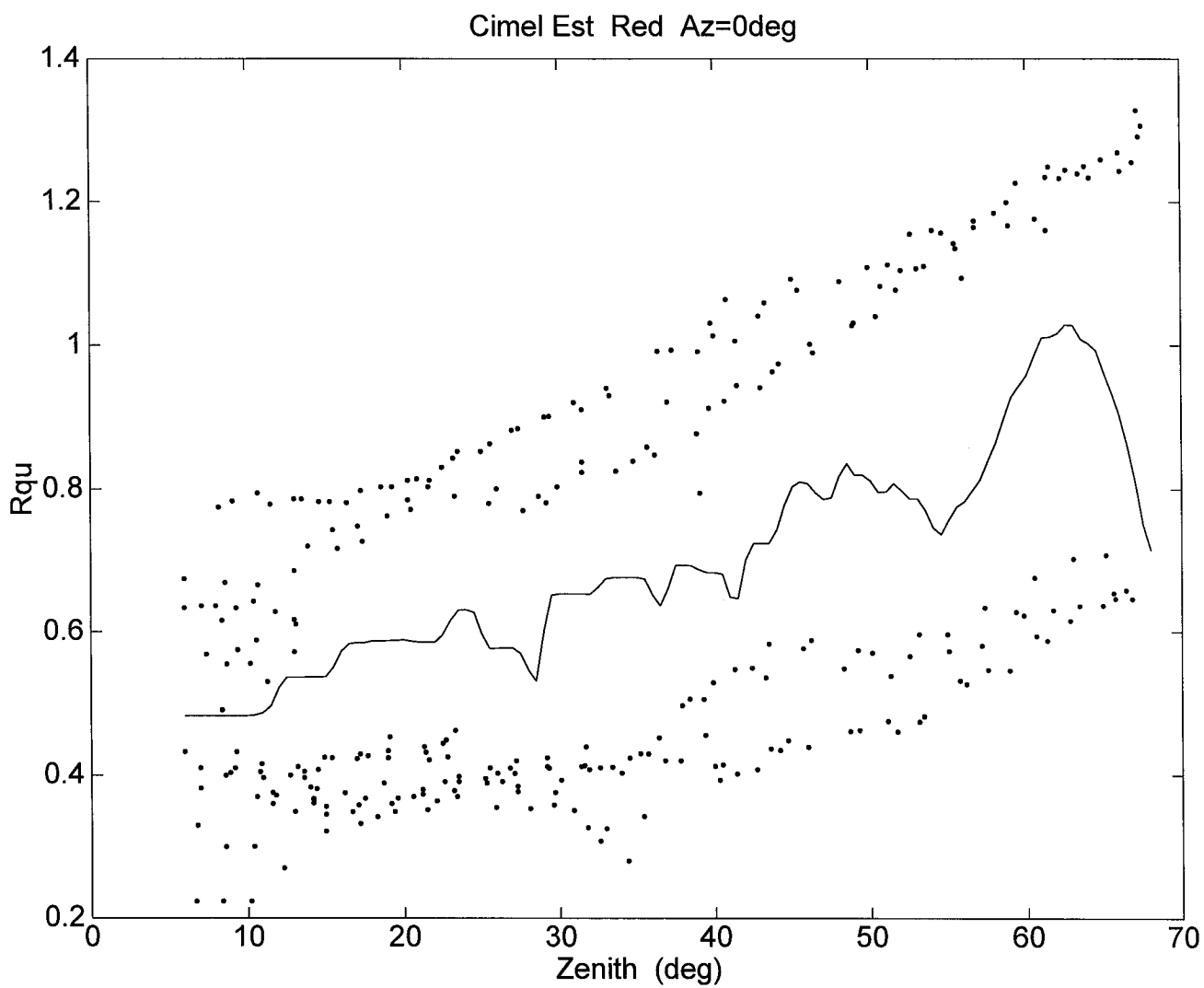
5 Appendix: Graphs of Results

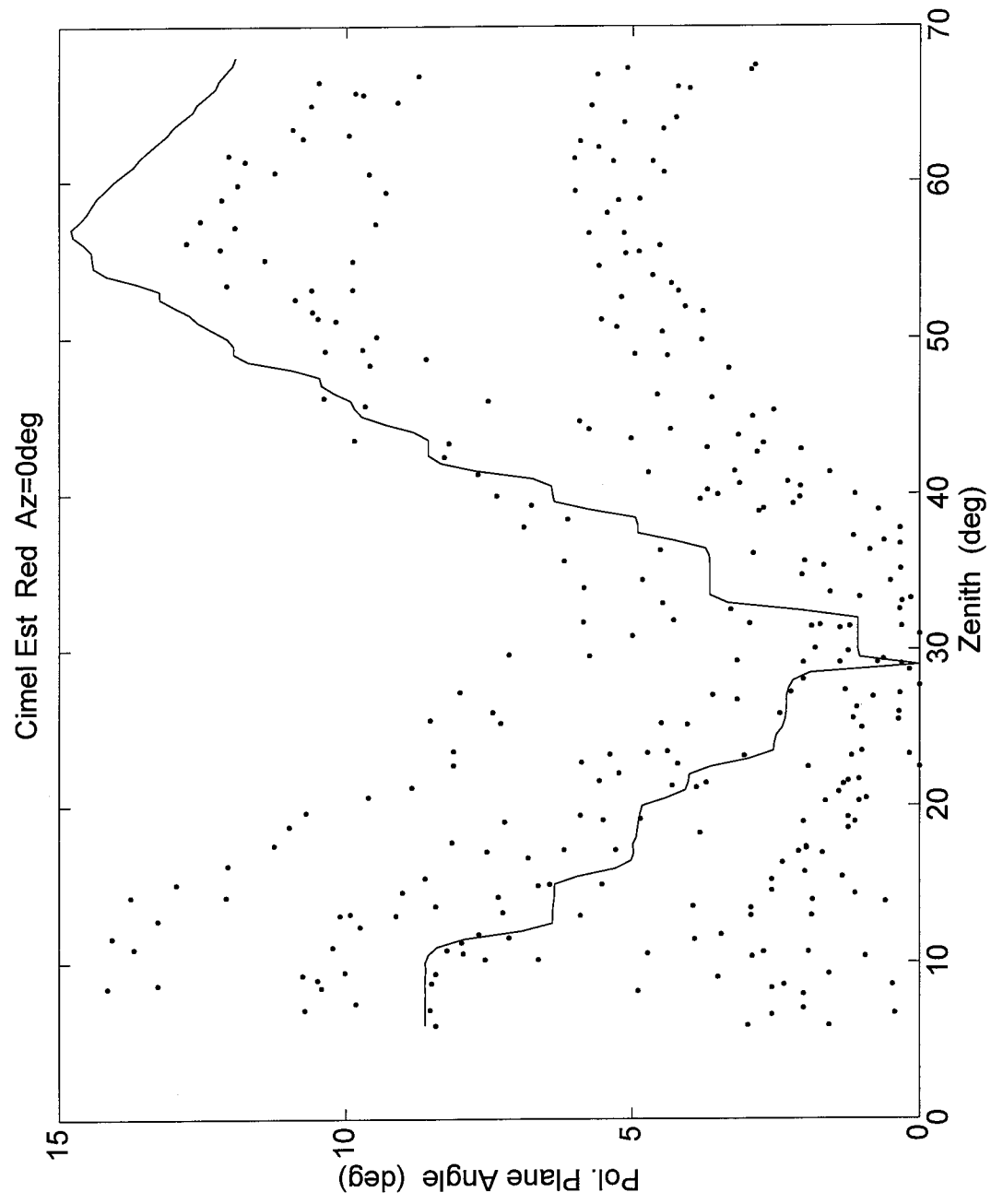
Key to linetypes:

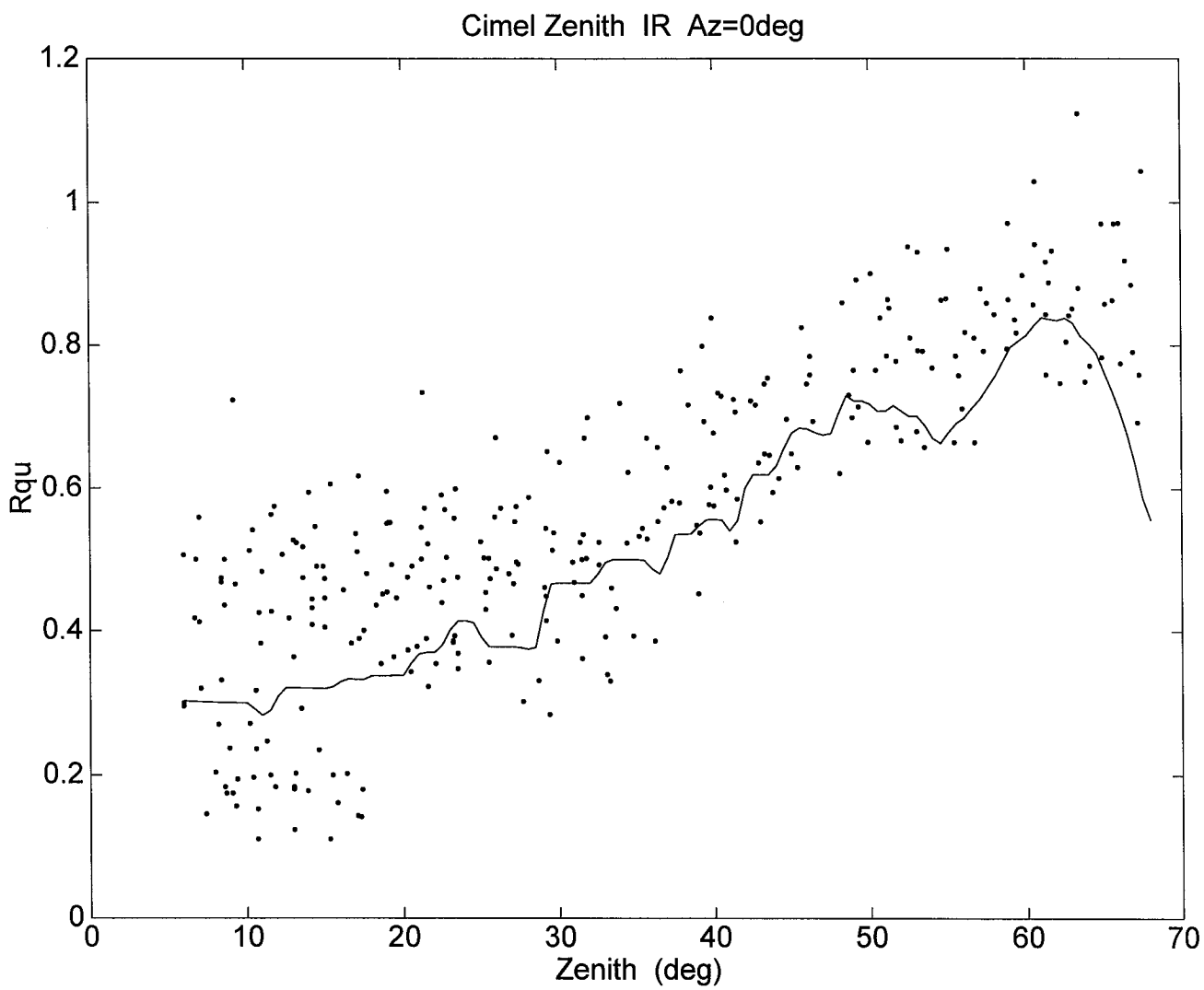
- dots: raw measurements.
- solid line: linear least-square solution; the a-priori information d_1 is taken to be 0.
- dashed line: solution based on the nonlinear cost function R_{QU} ; the a-priori information $\hat{\rho}$ is taken to be the mean R_{QU} corresponding to the linear least-square solution d_0 for the specified view azimuth.

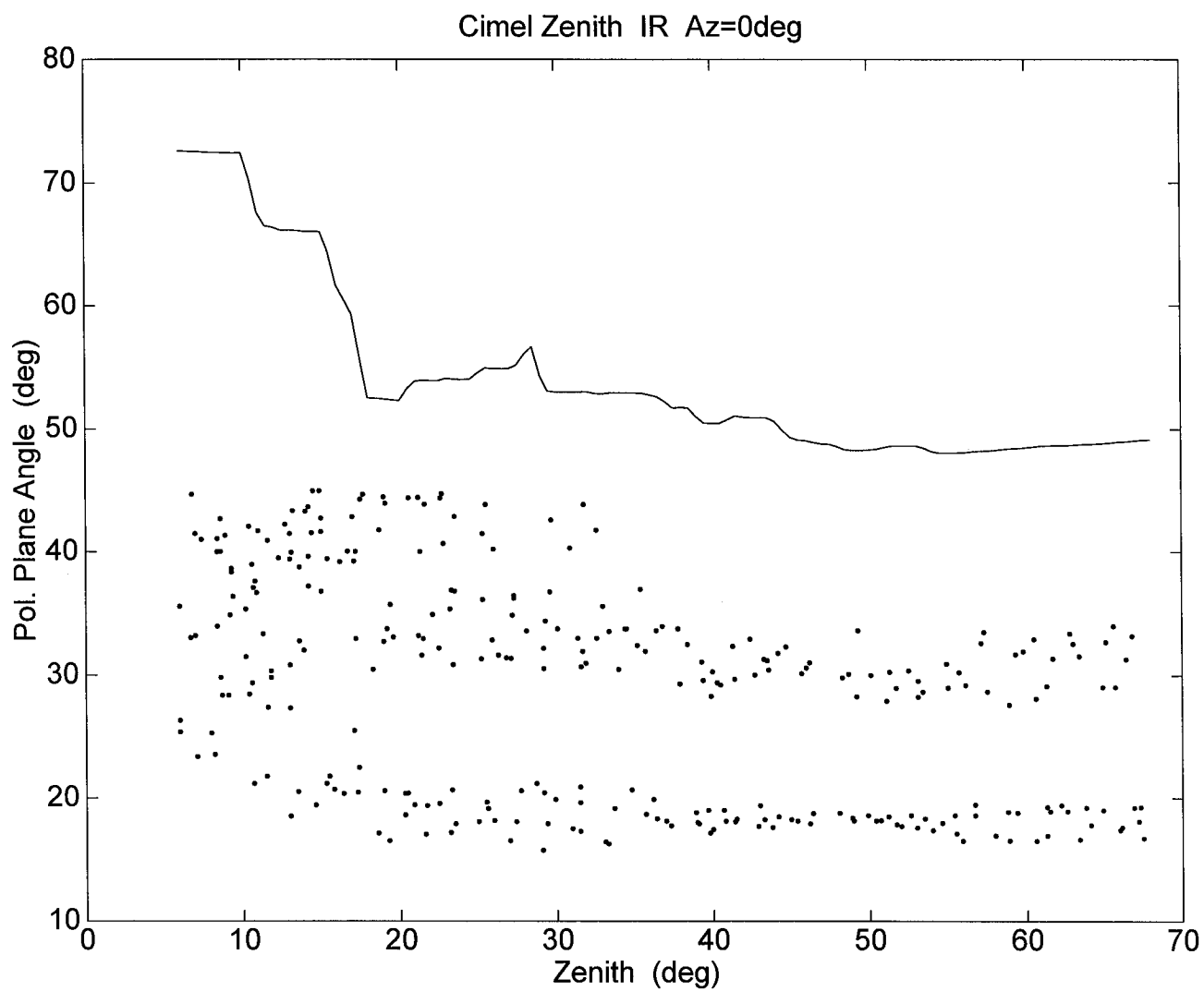
All graphs are for measurements taken on 07/25/91.

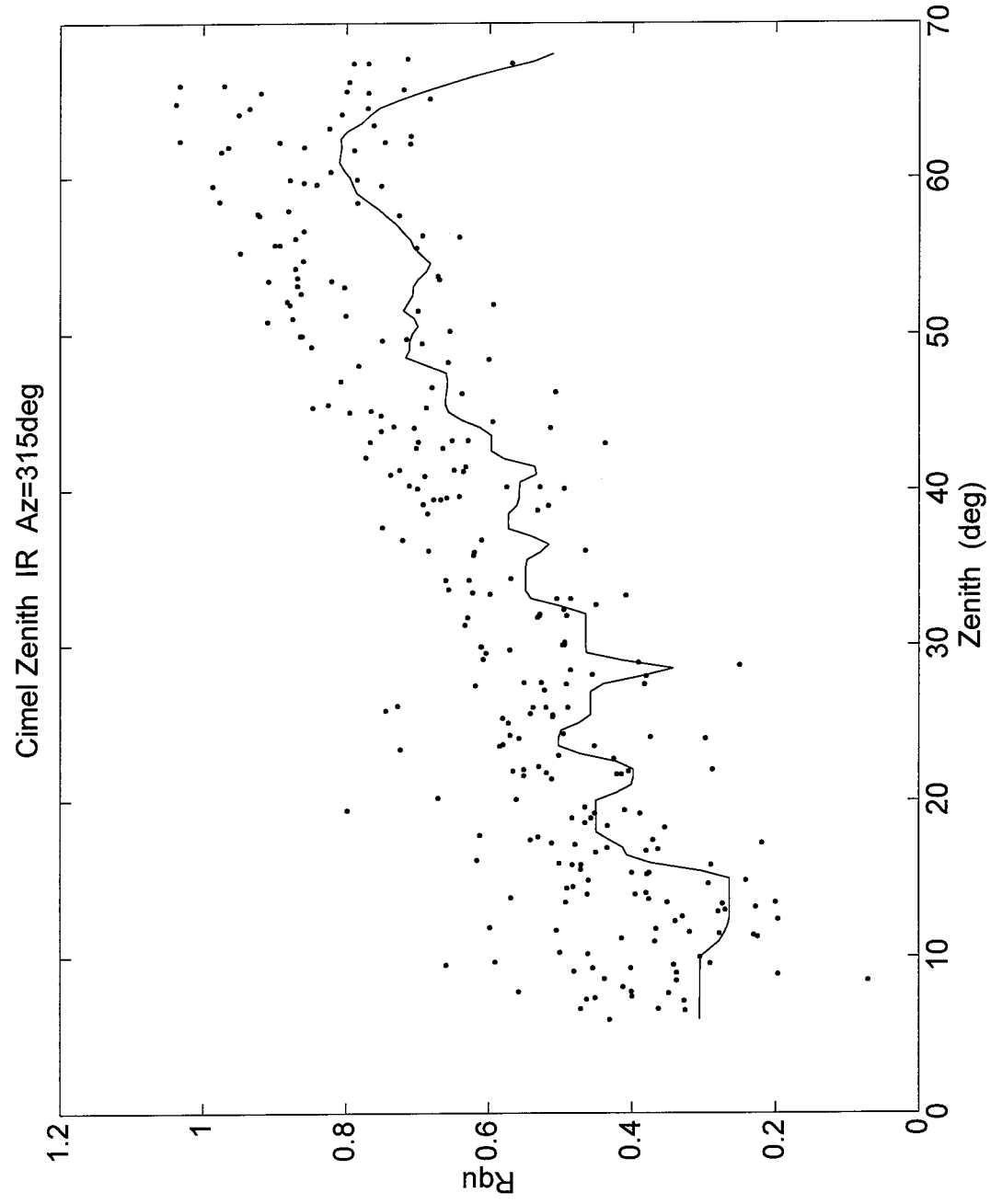
Barnes instruments had raw FOV of 1° and Cimel instruments had raw FOV of 12° , with measurements taken over irregular azimuth and zenith view angles. The processed data, by either method, used an effective ('simulated') FOV of 12° at regularly spaced zenith angles for fixed azimuth angles taken from 0° to 315° in increments of 45° . Not all results are shown.

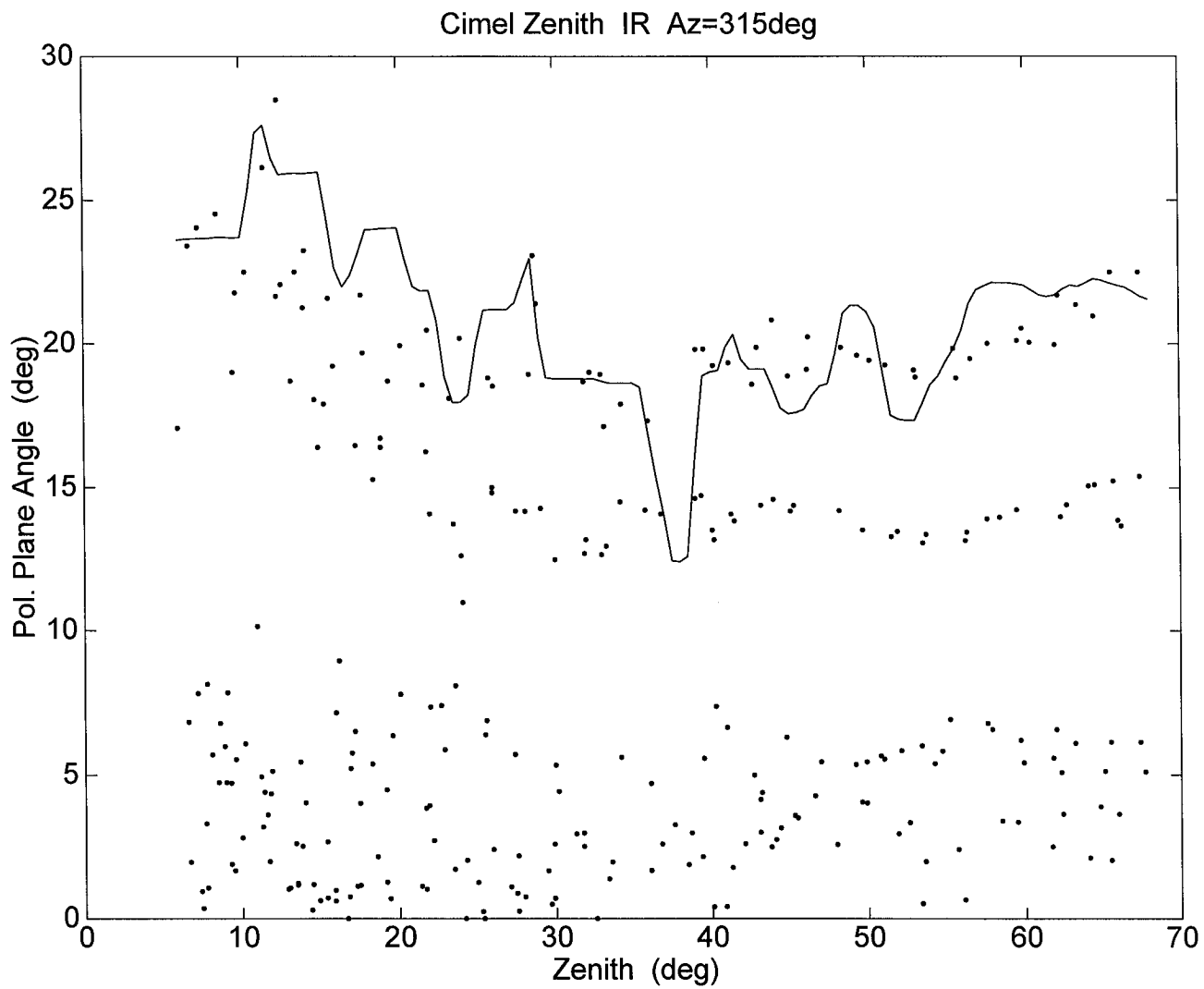


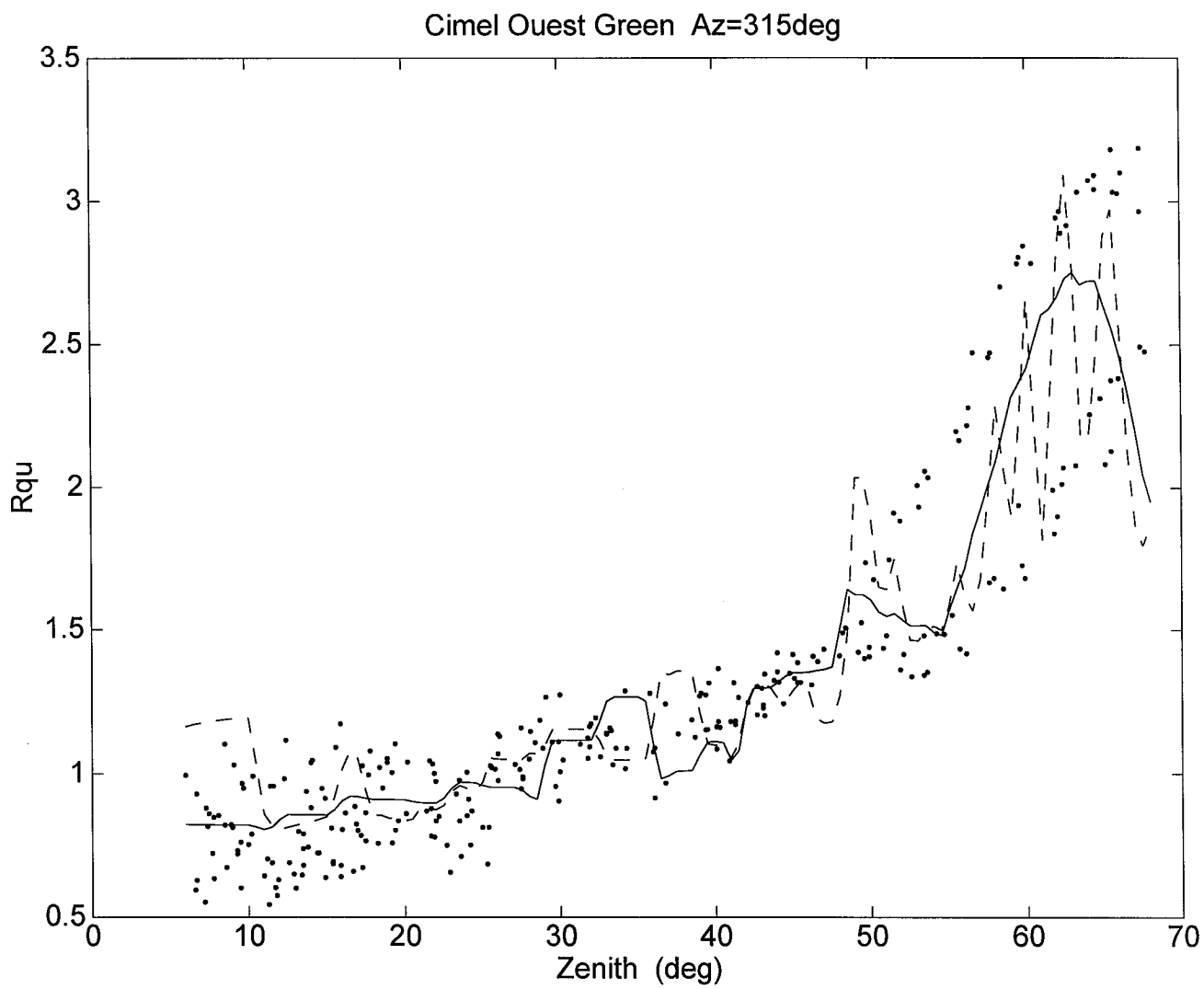


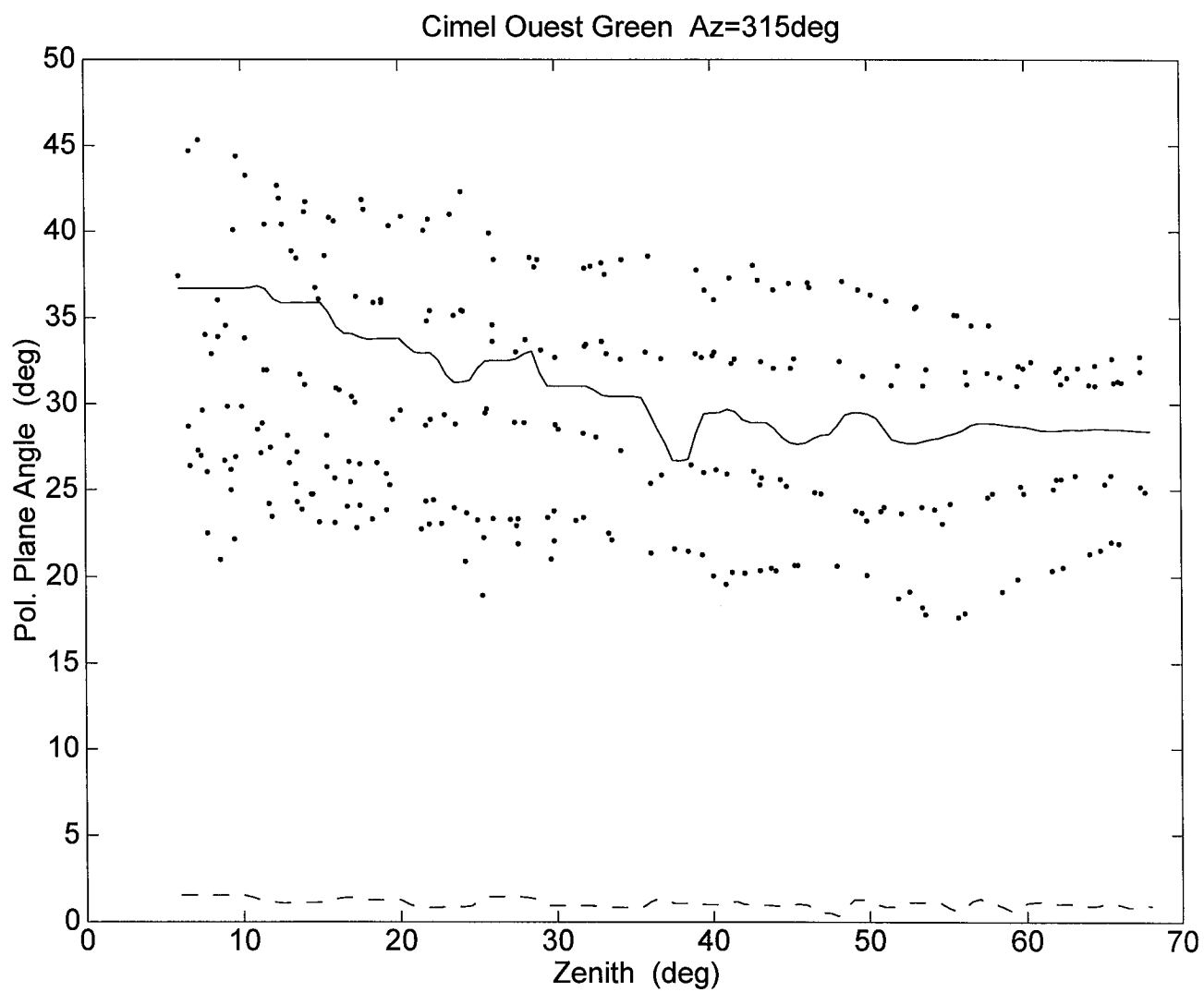


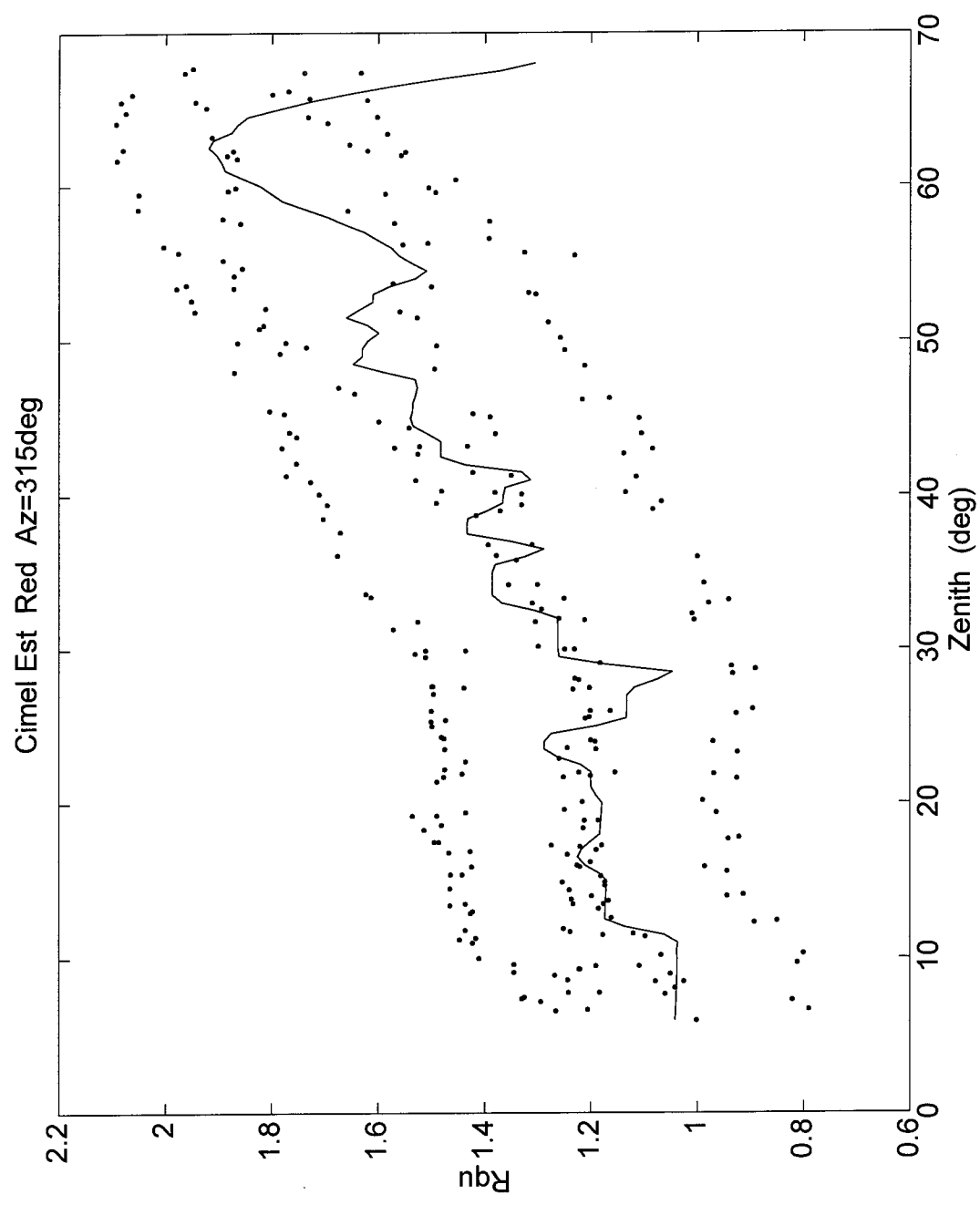




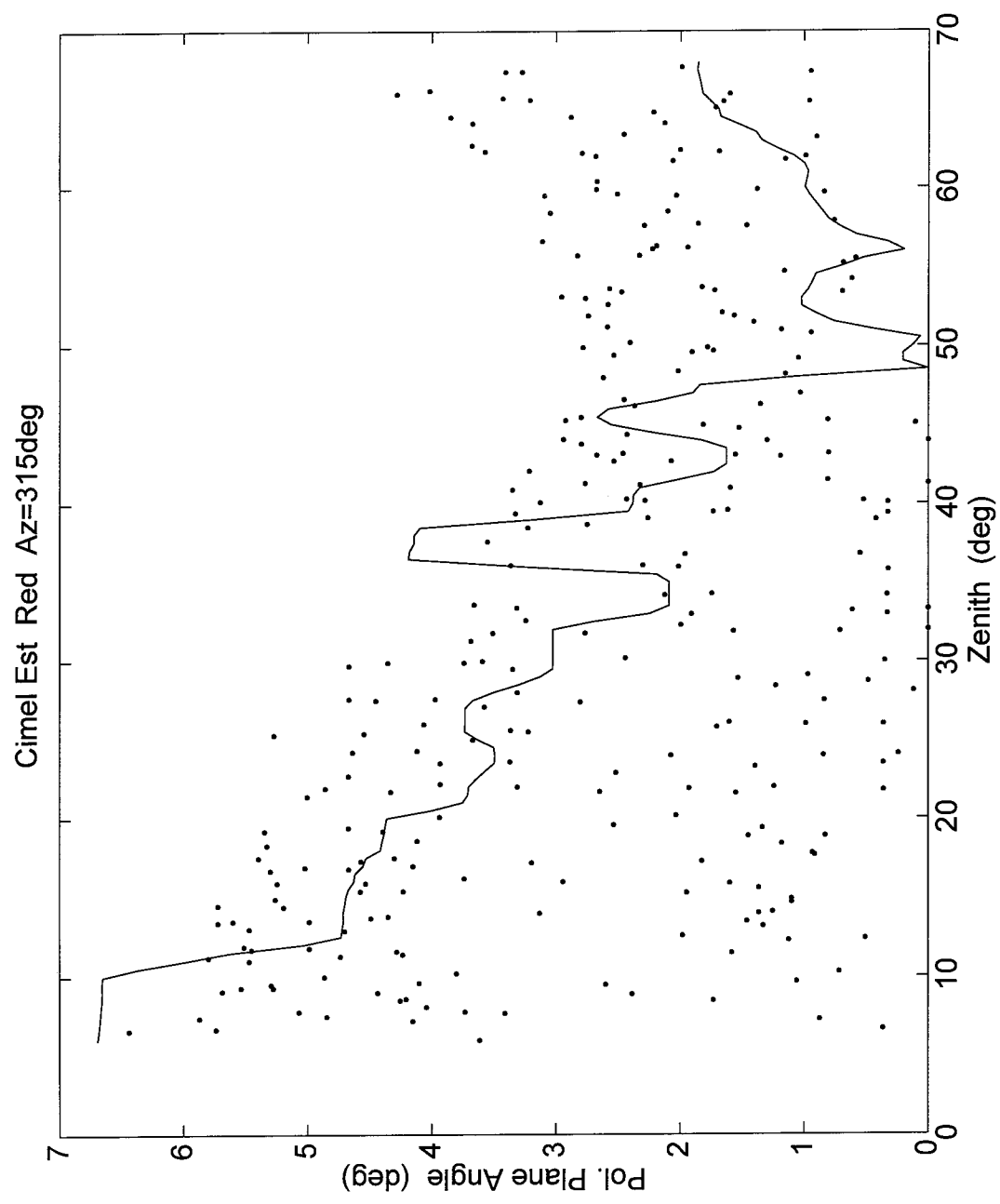


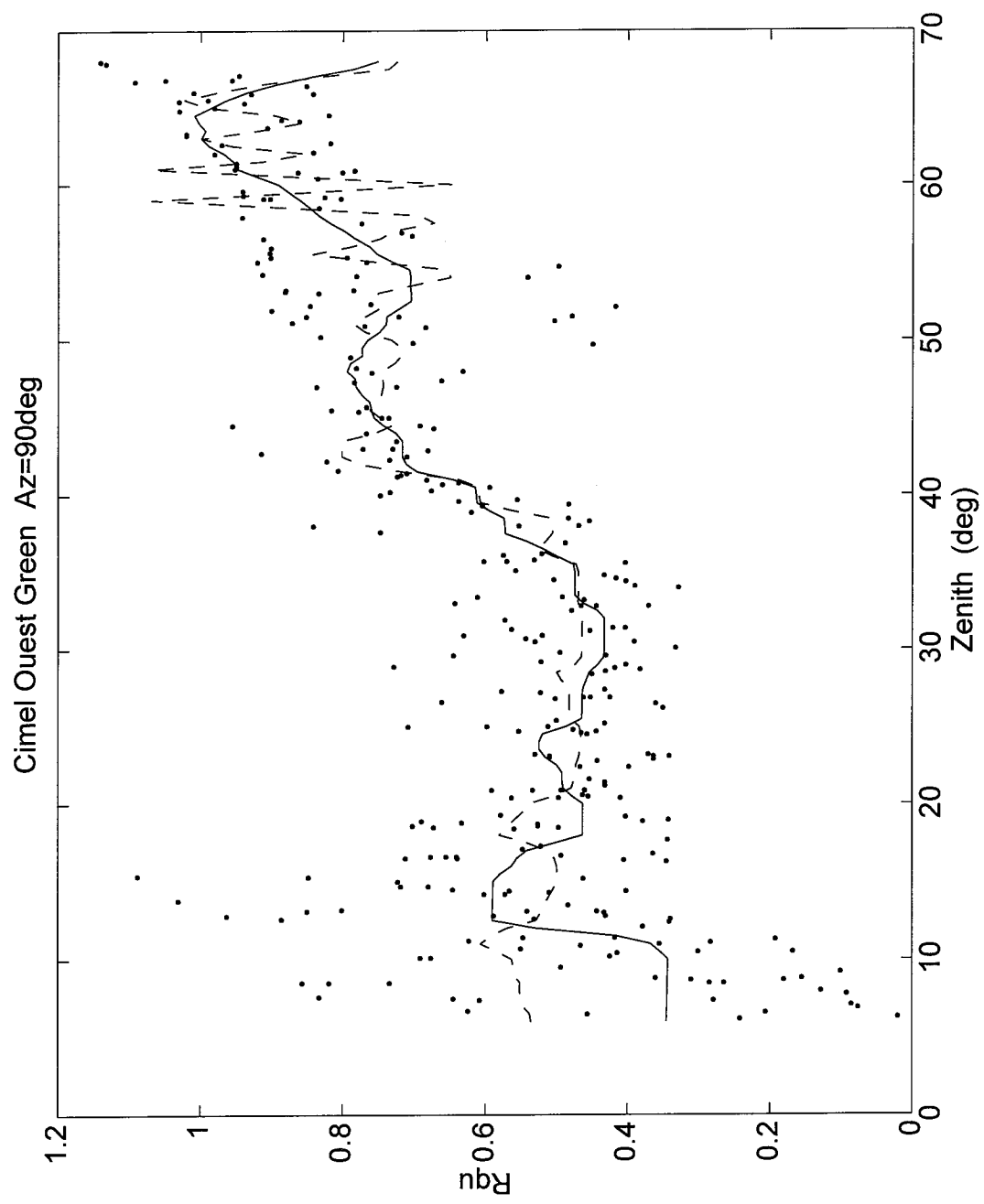


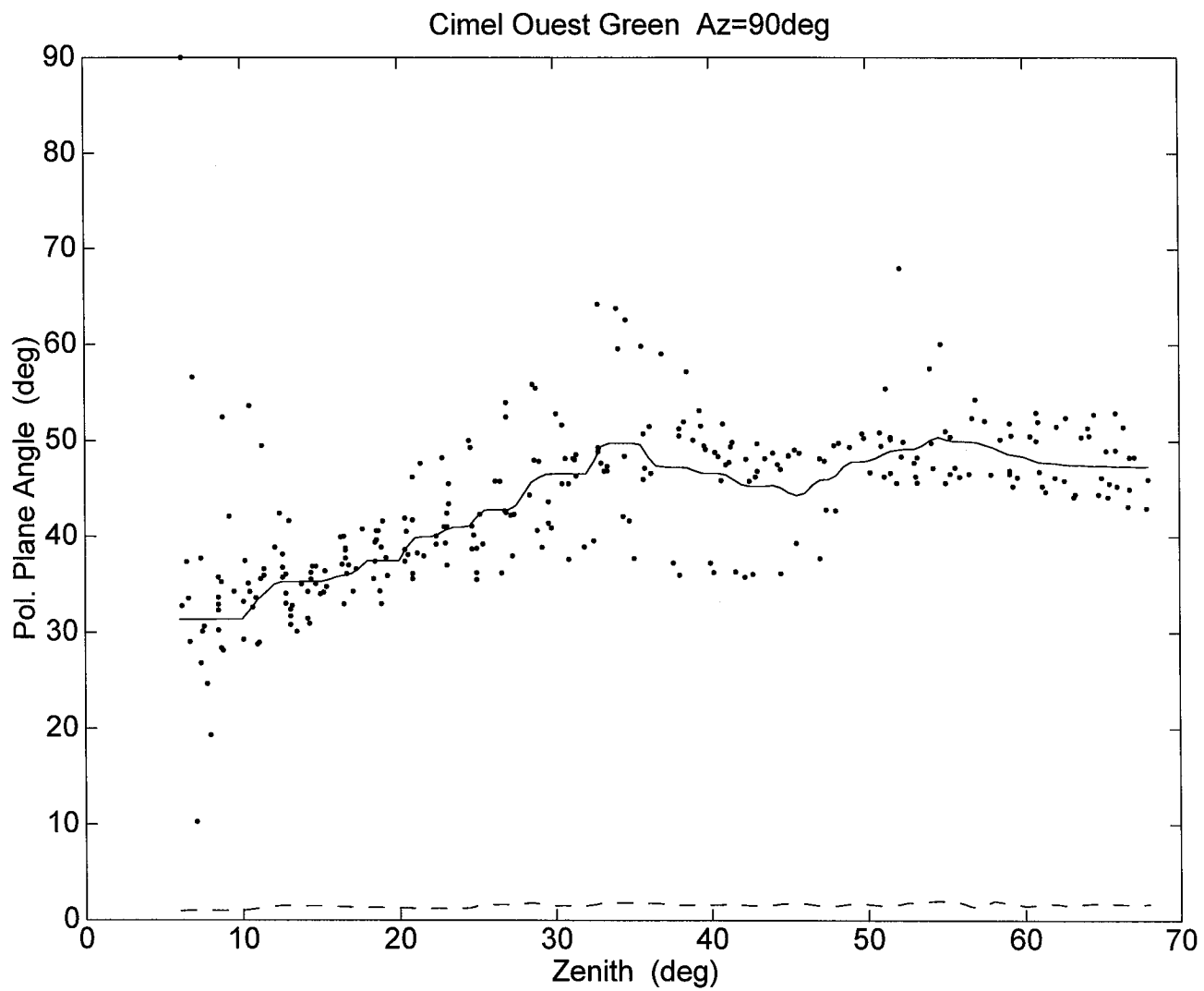


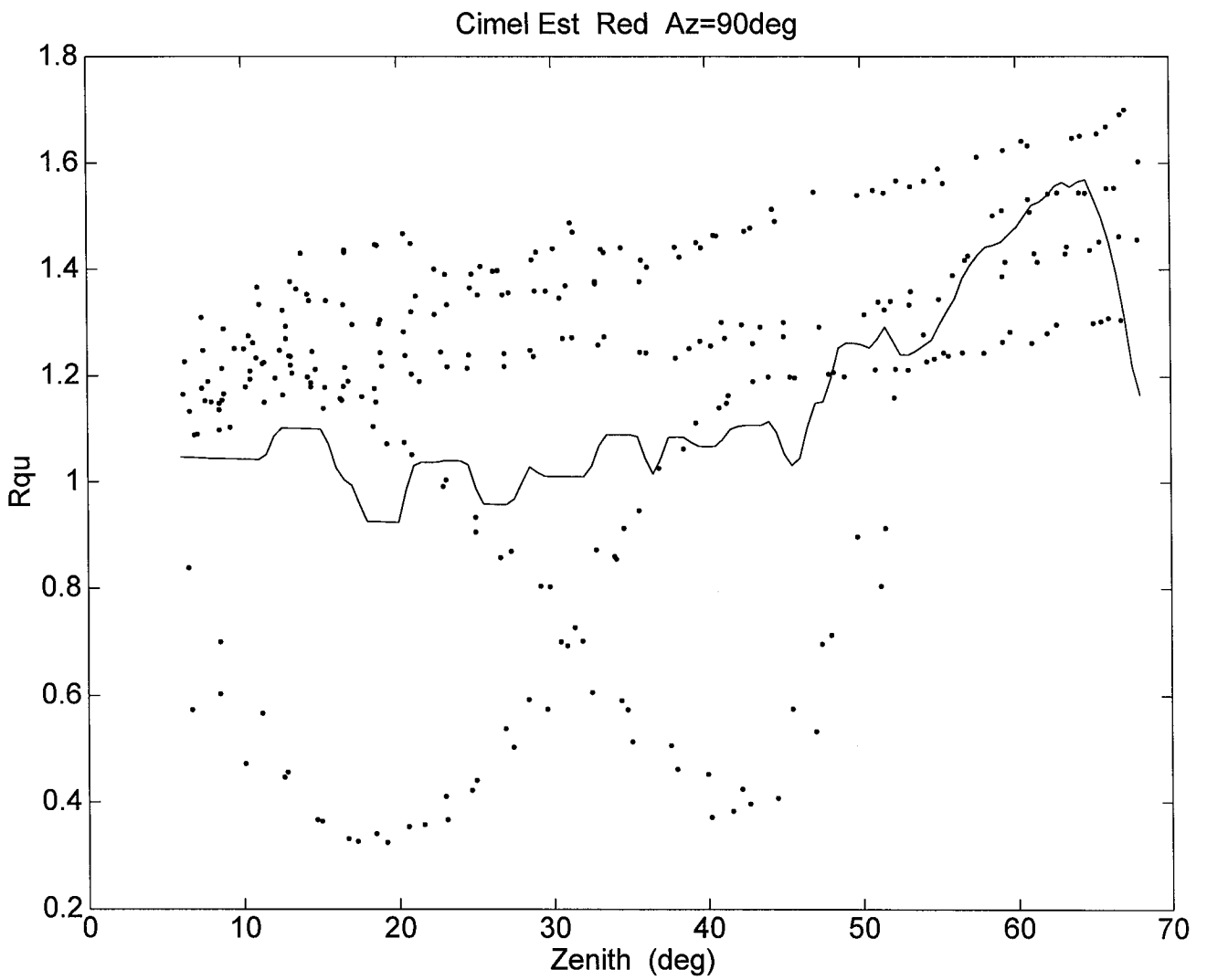


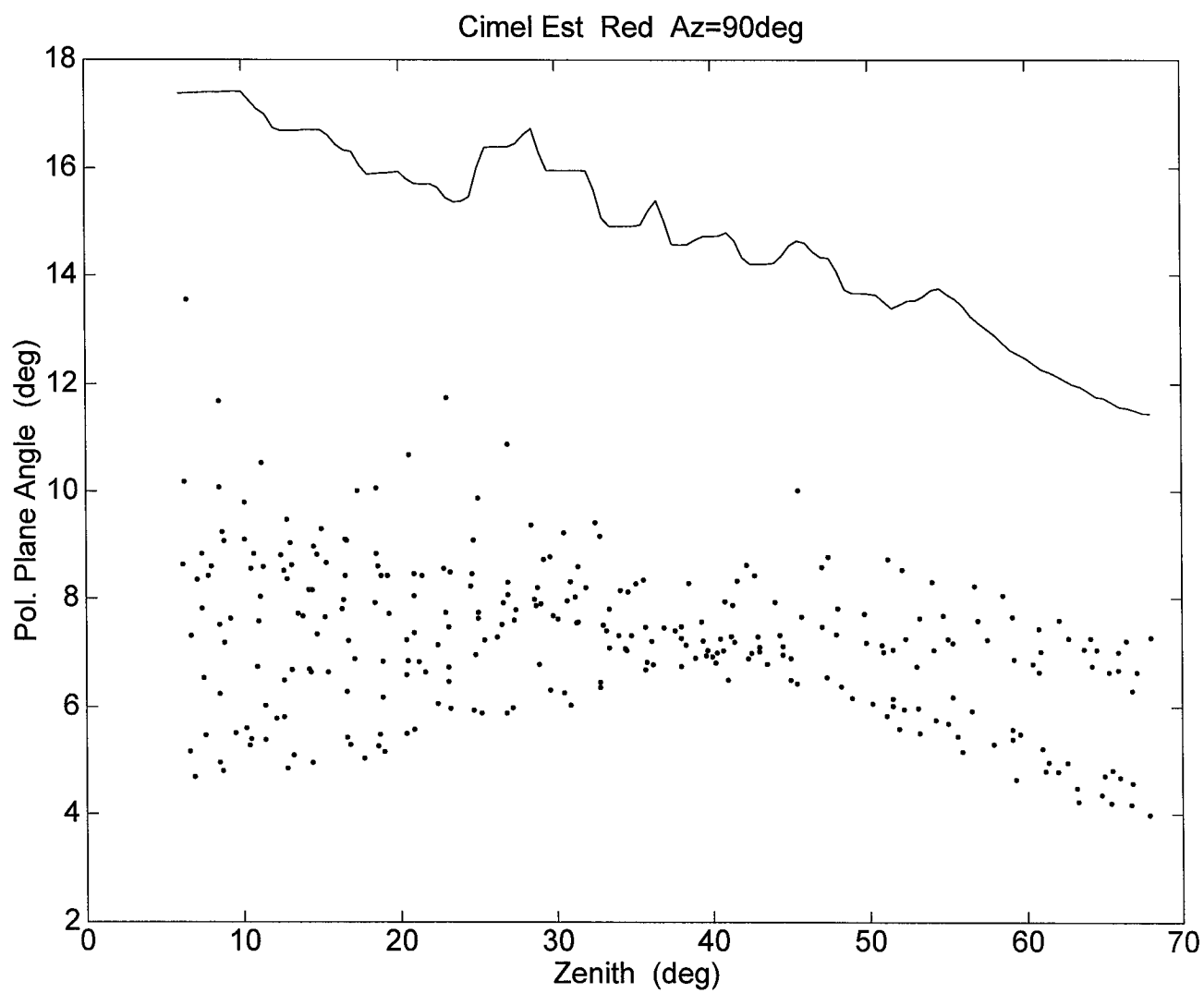
LL

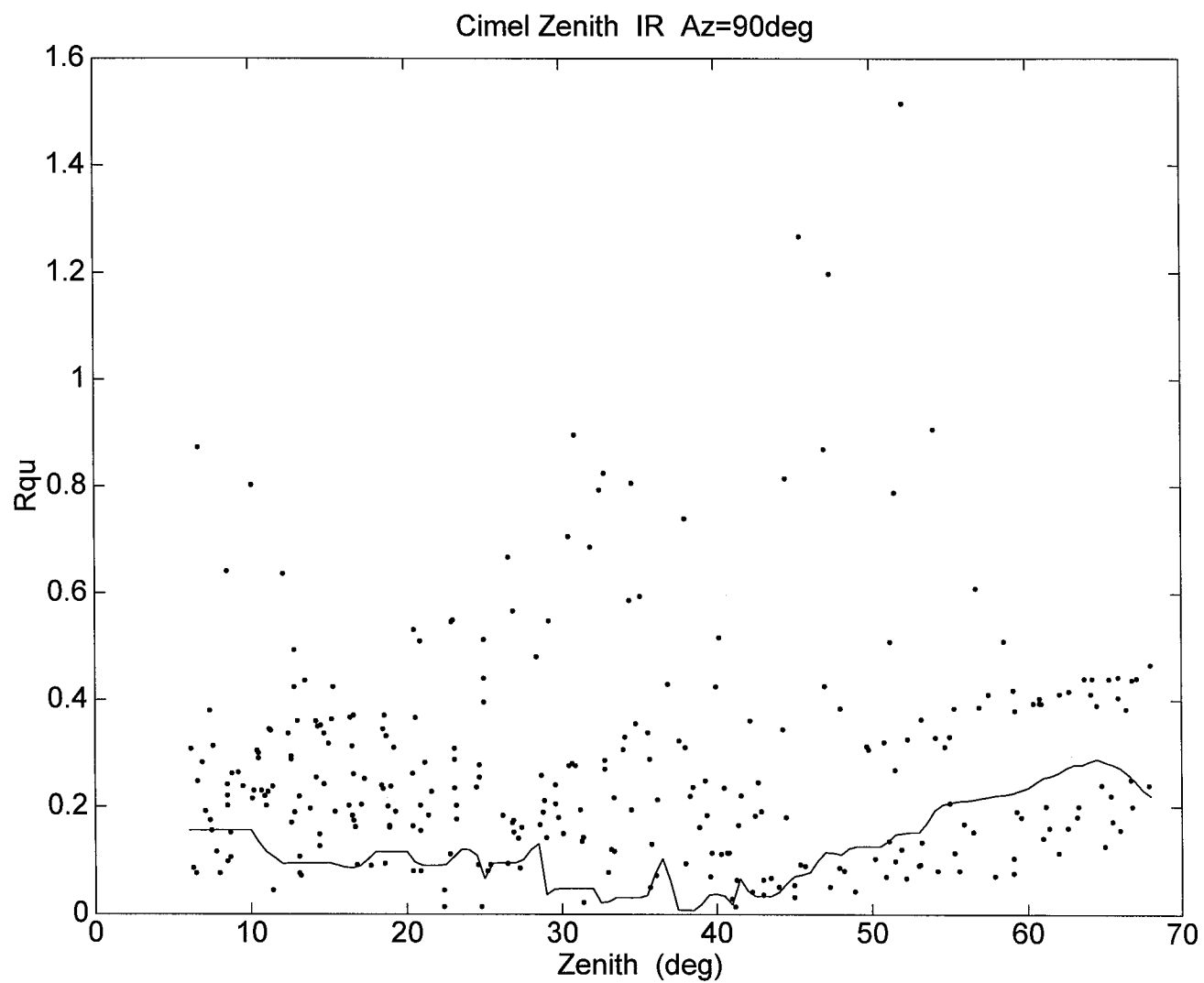


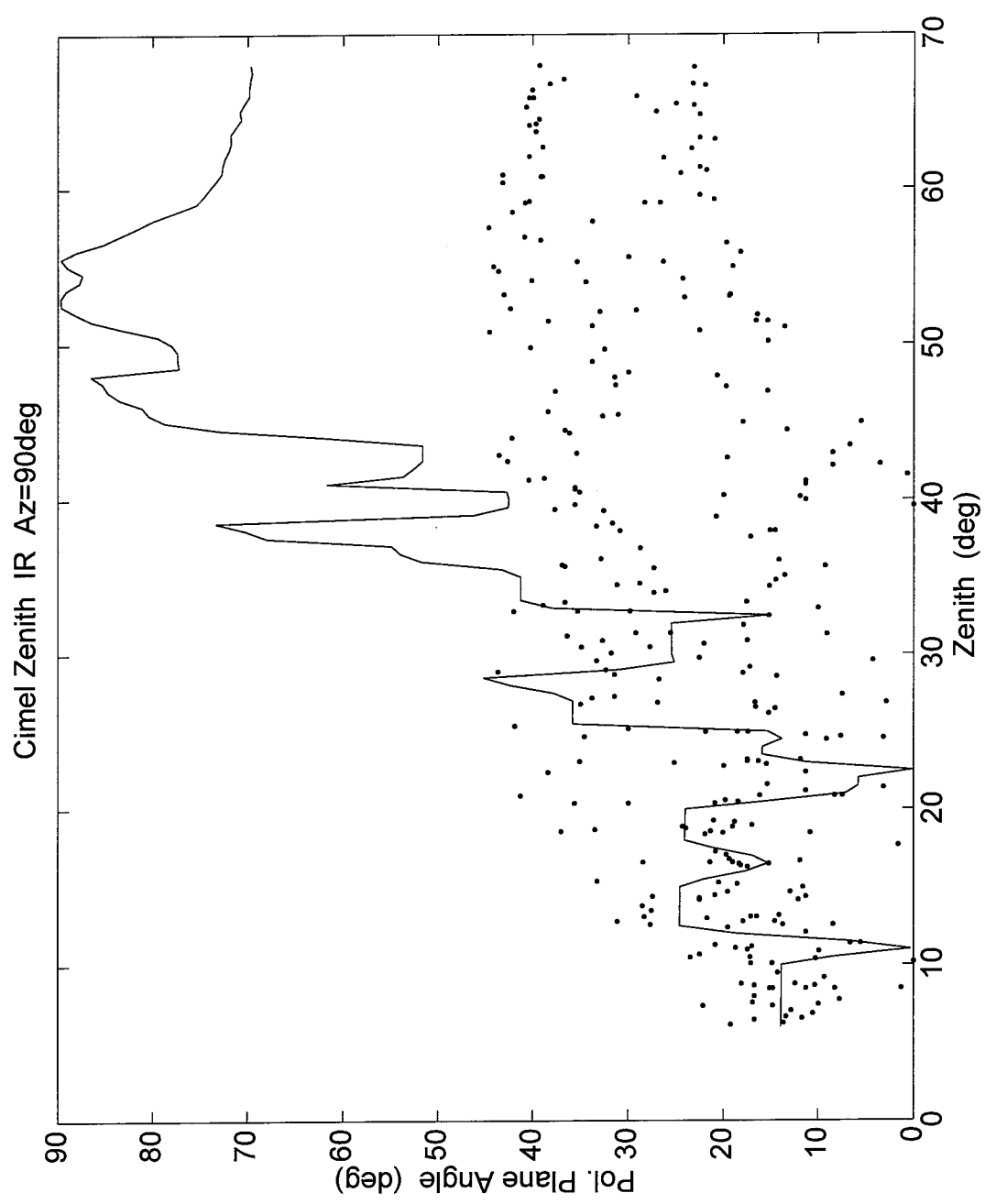


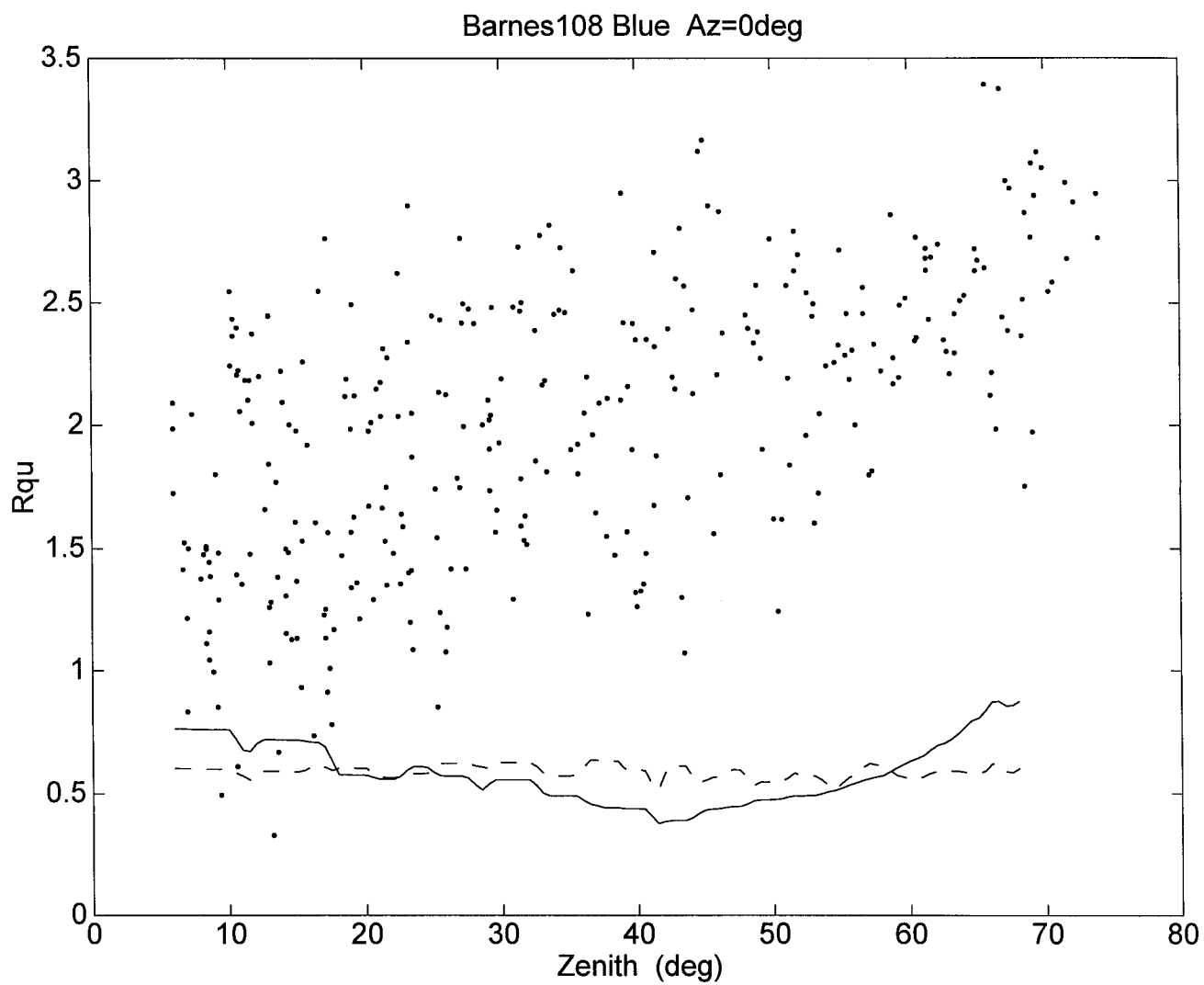


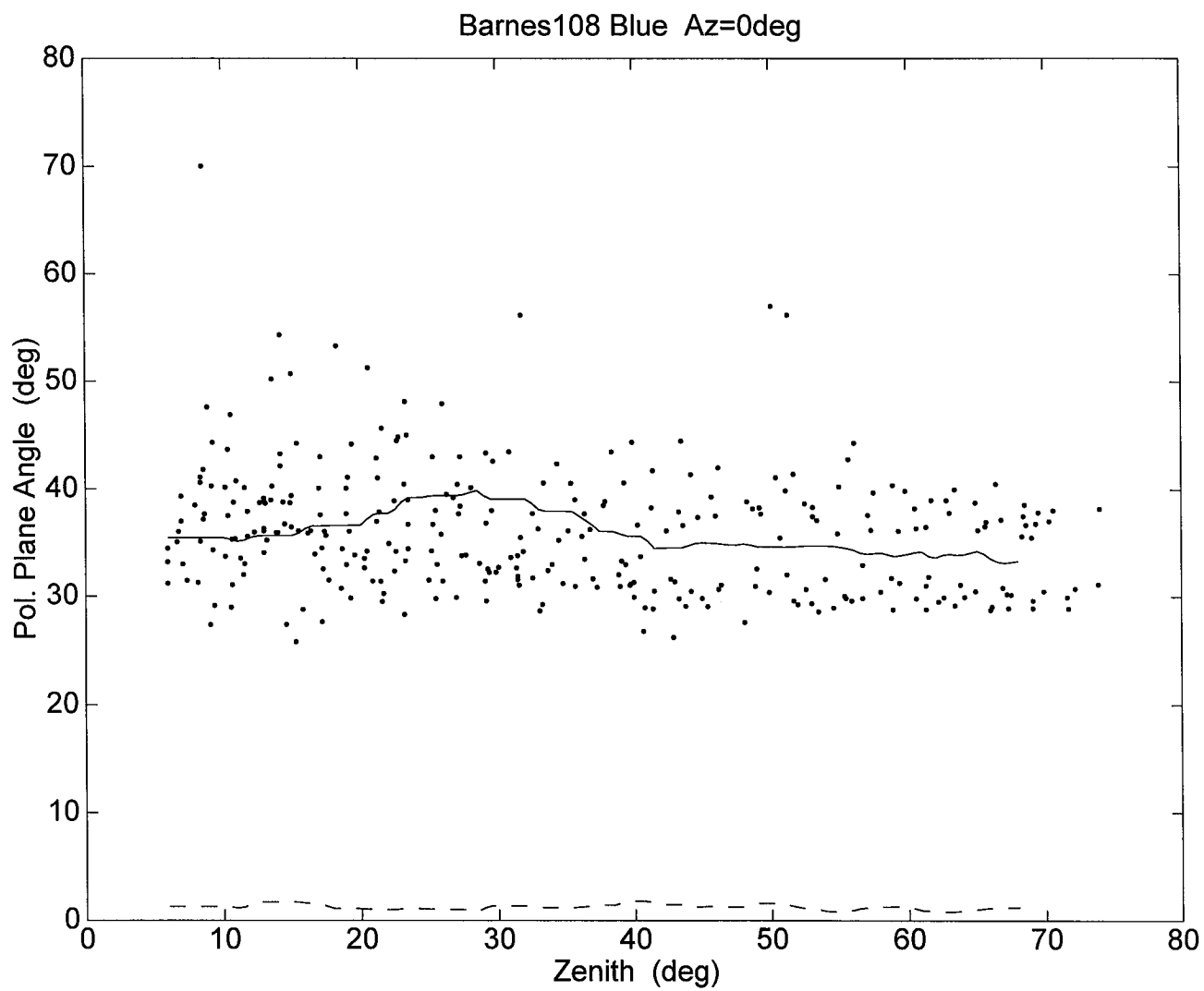


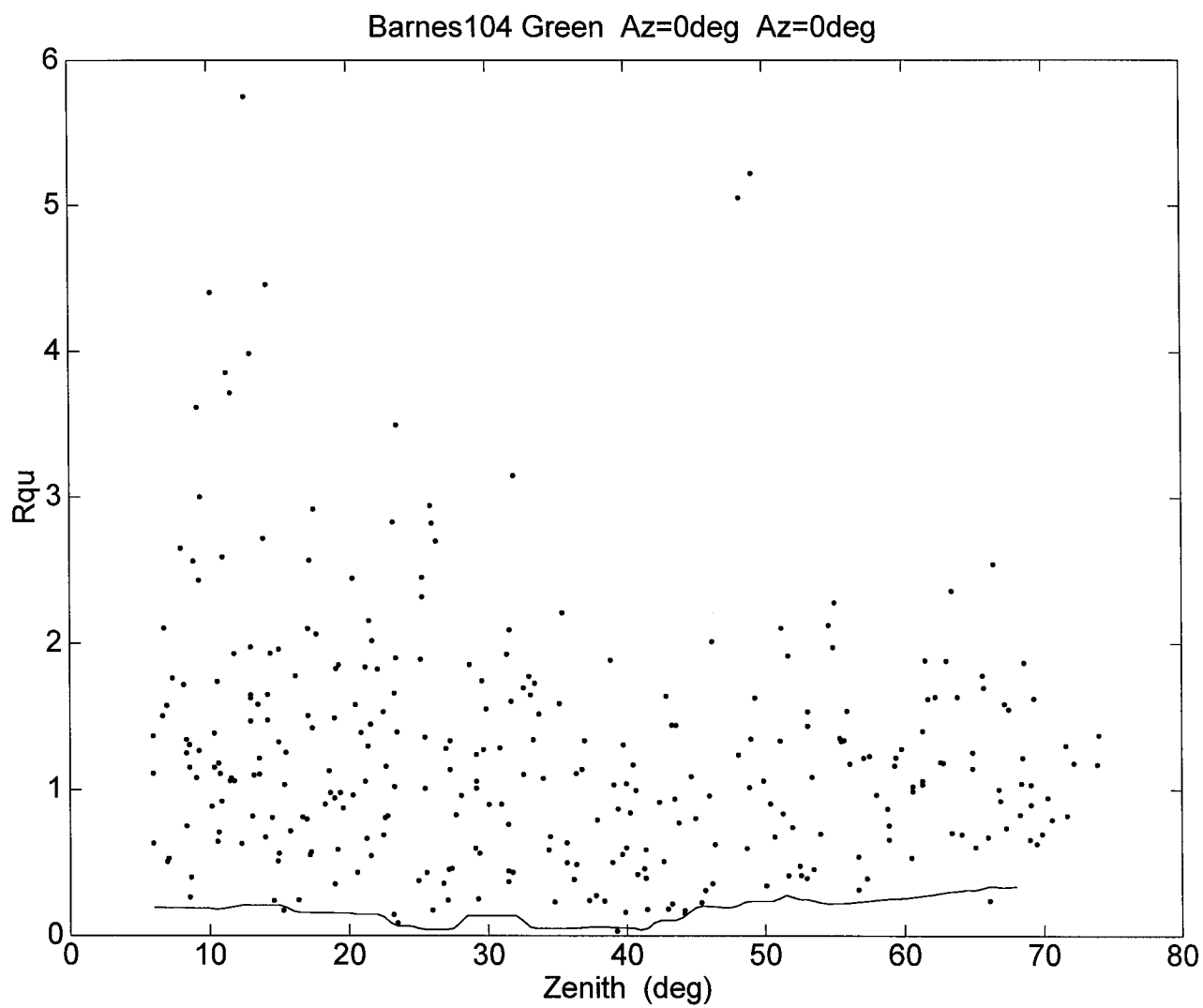


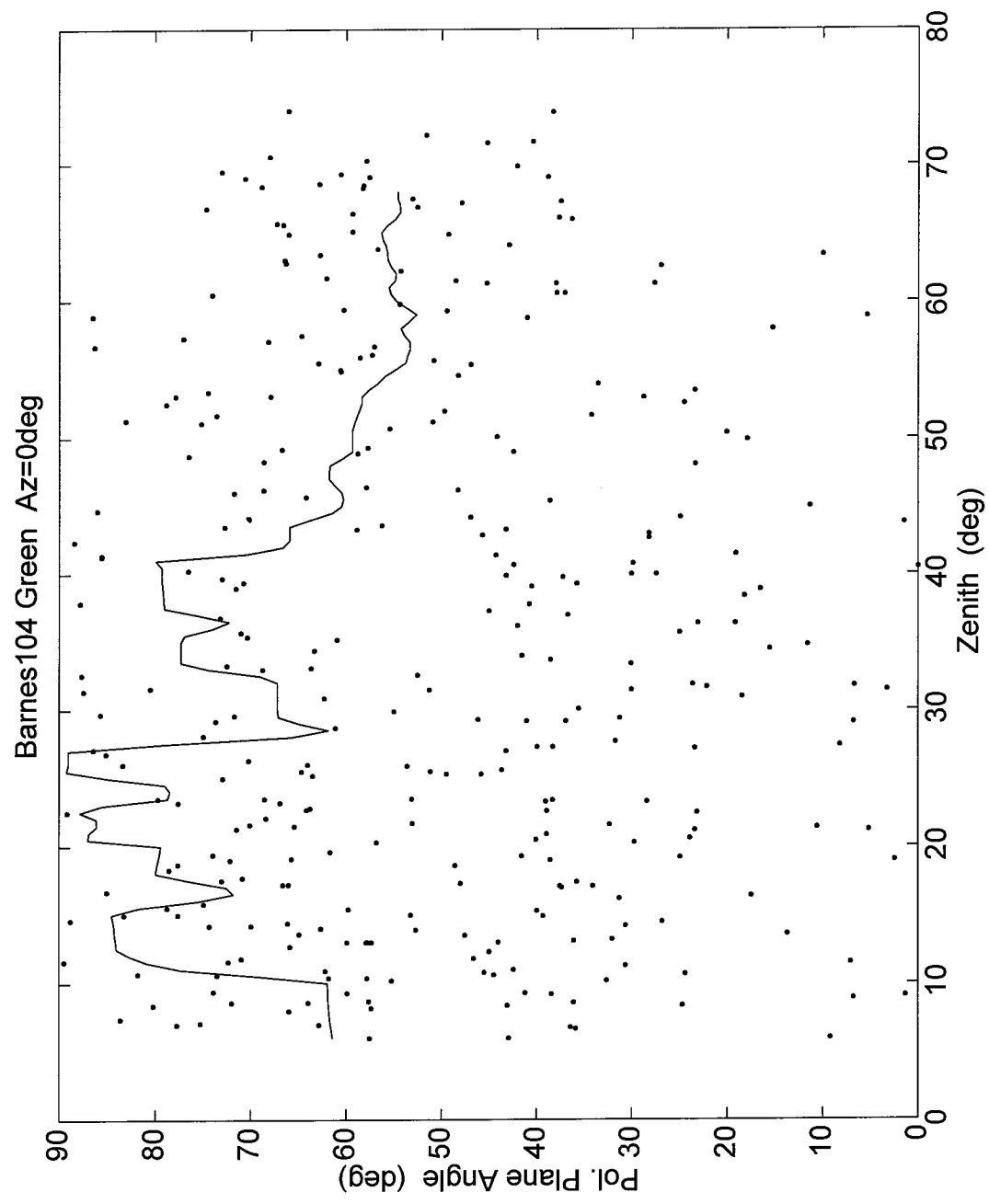


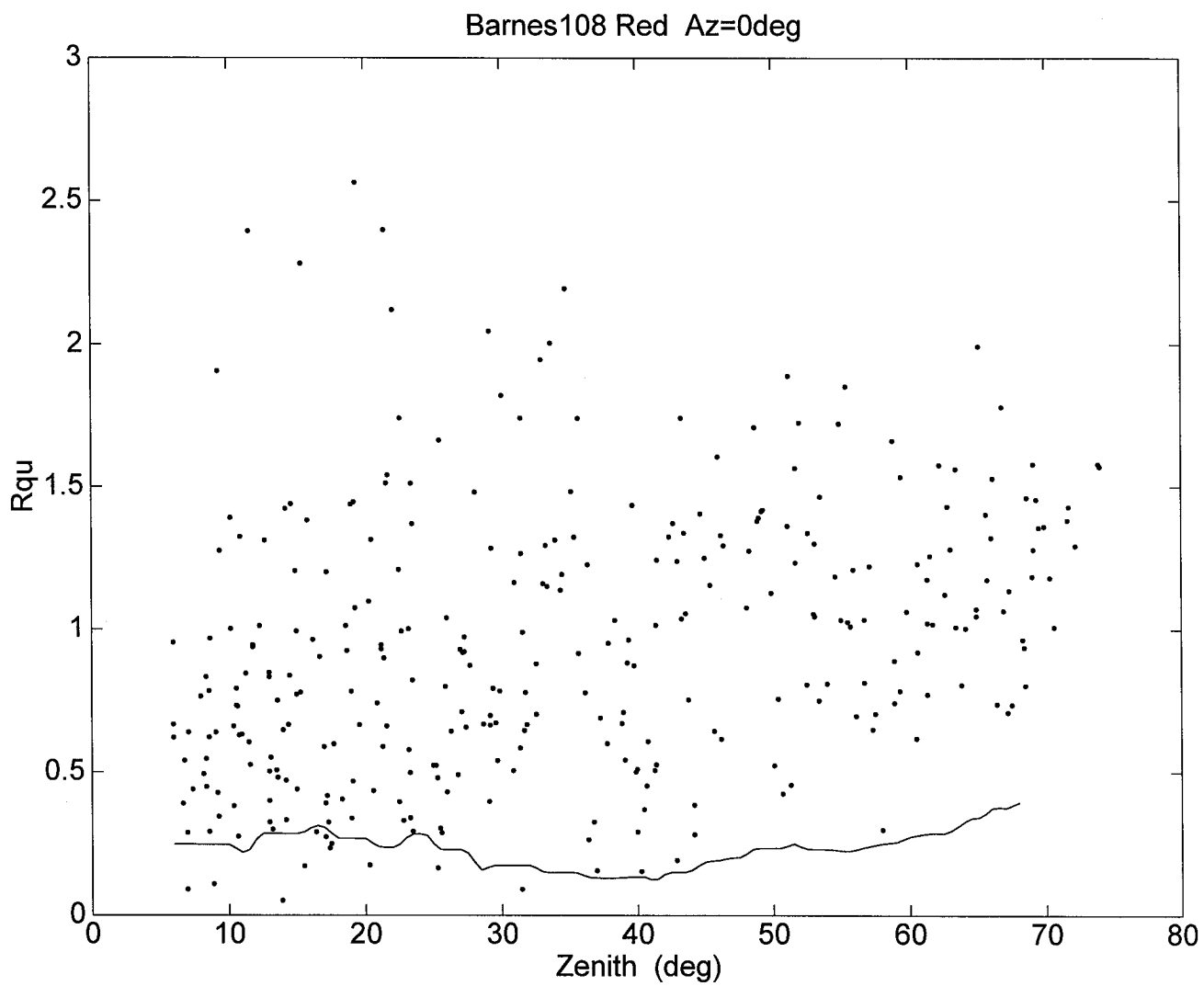


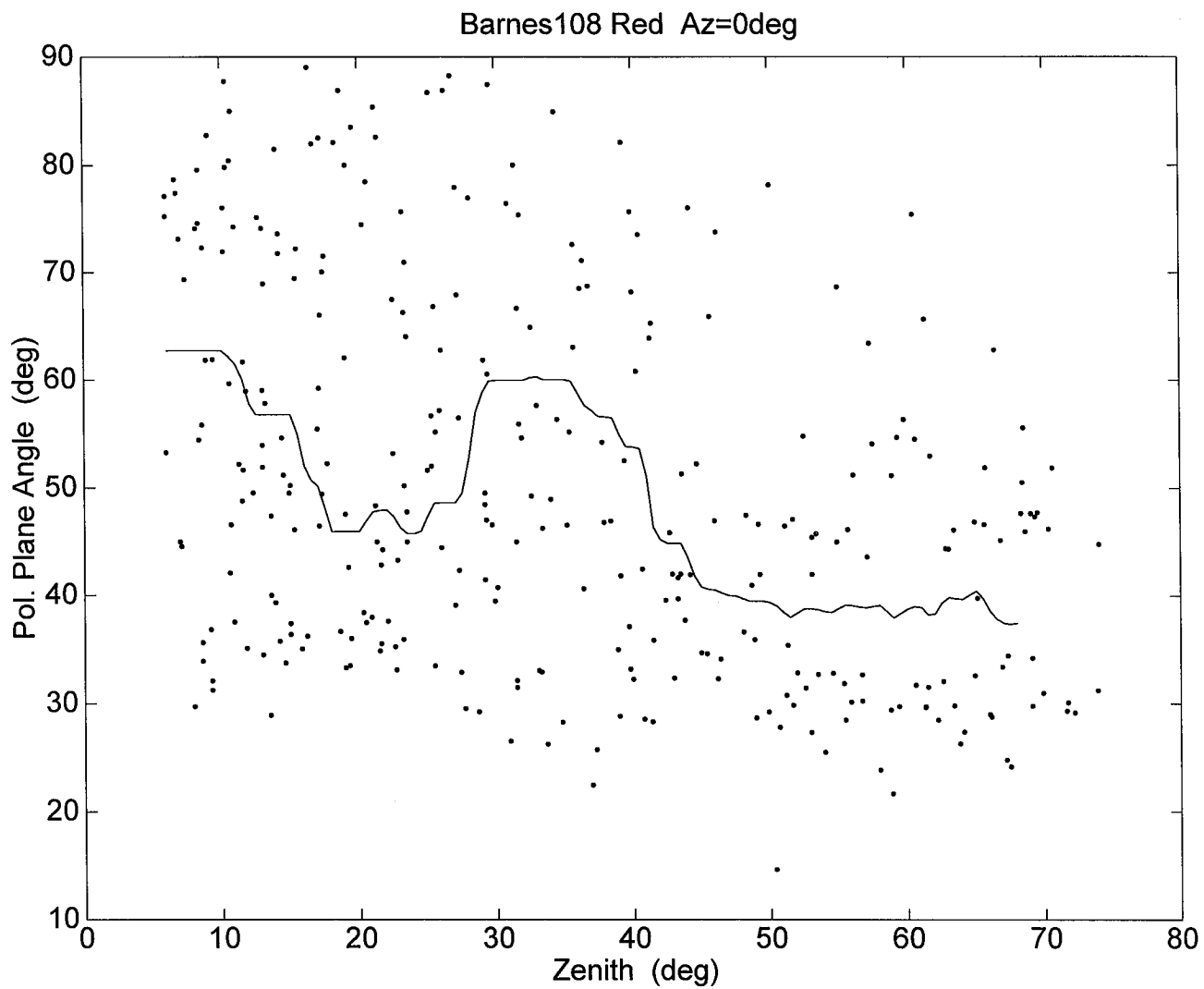


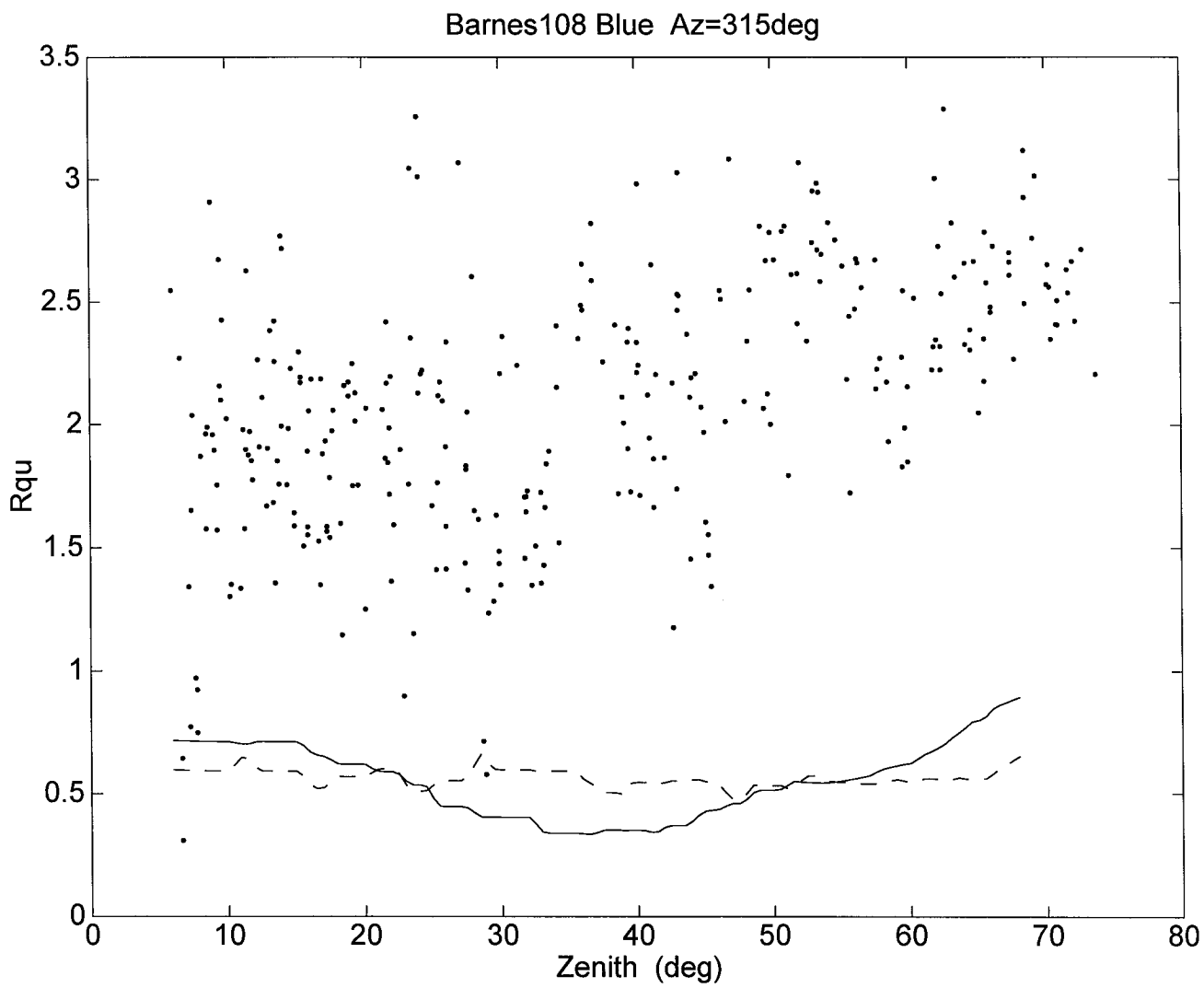


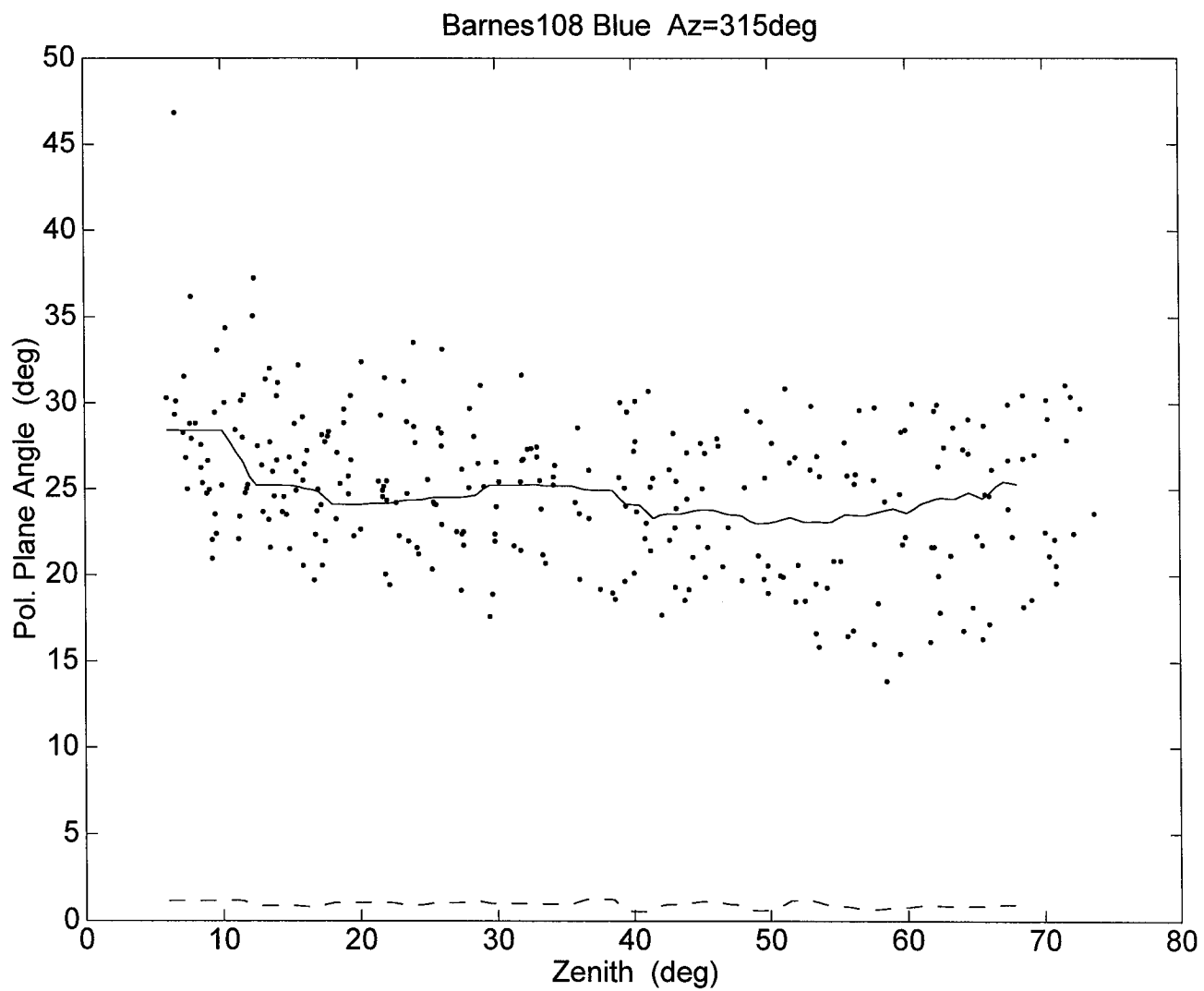


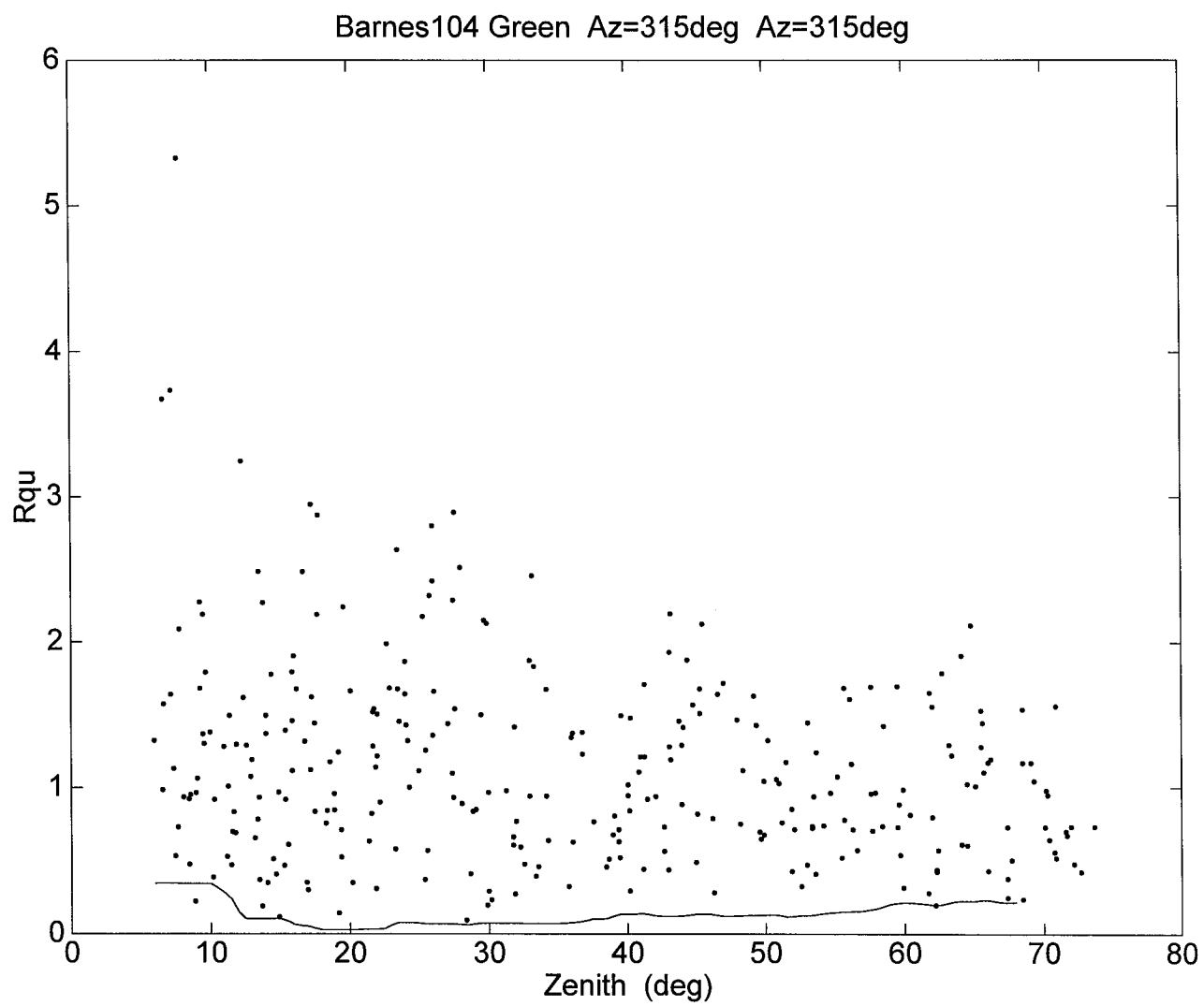


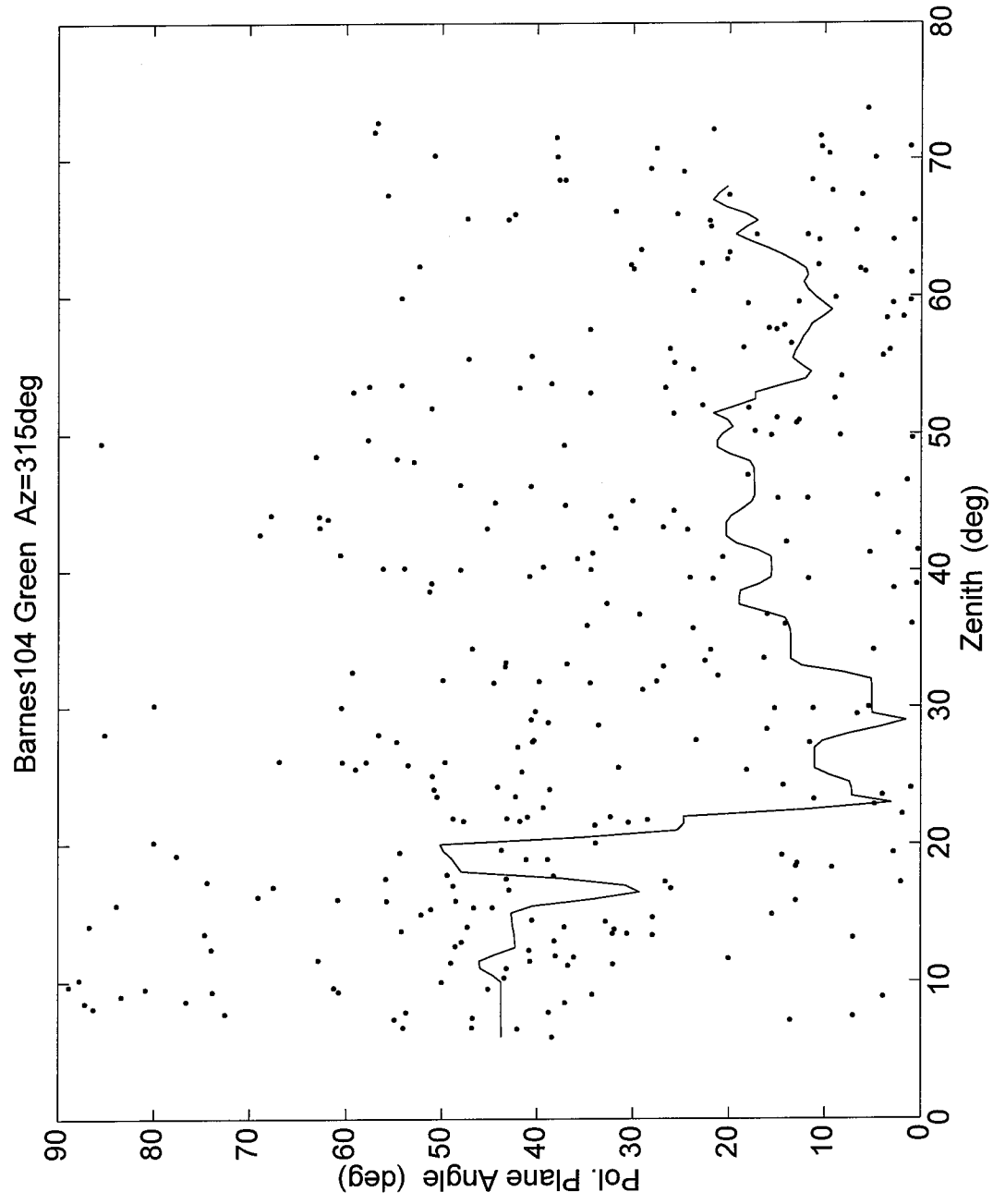


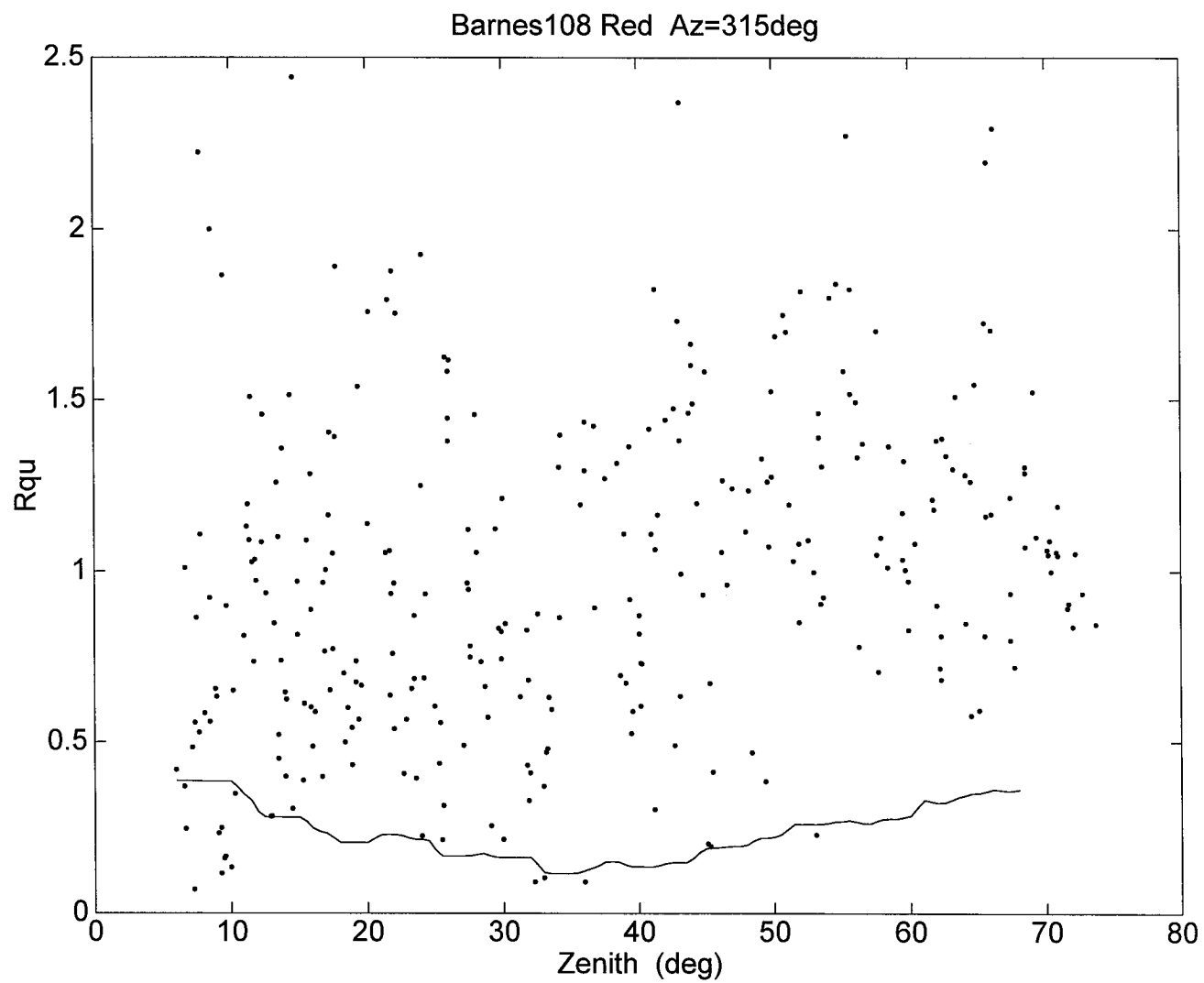


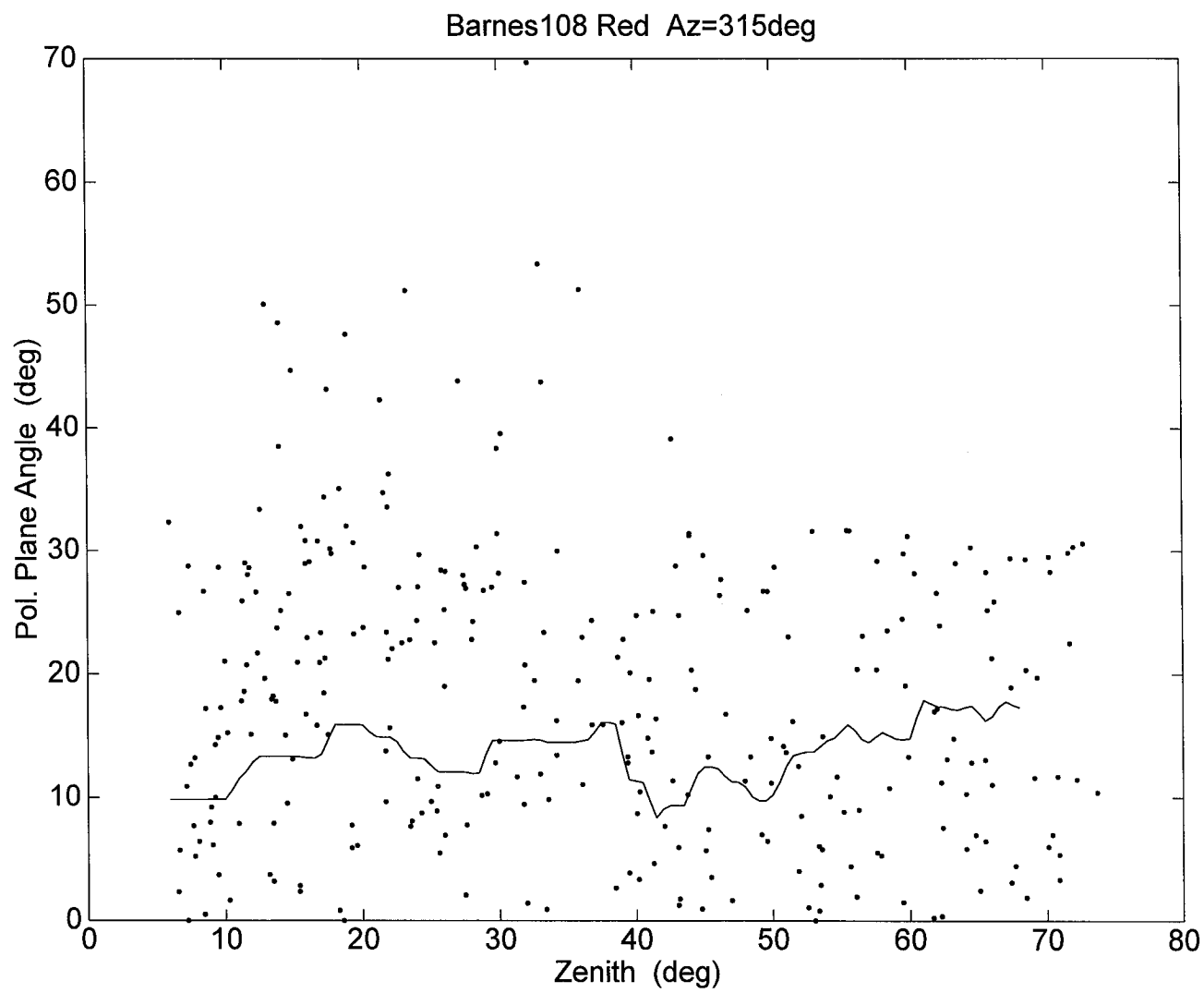


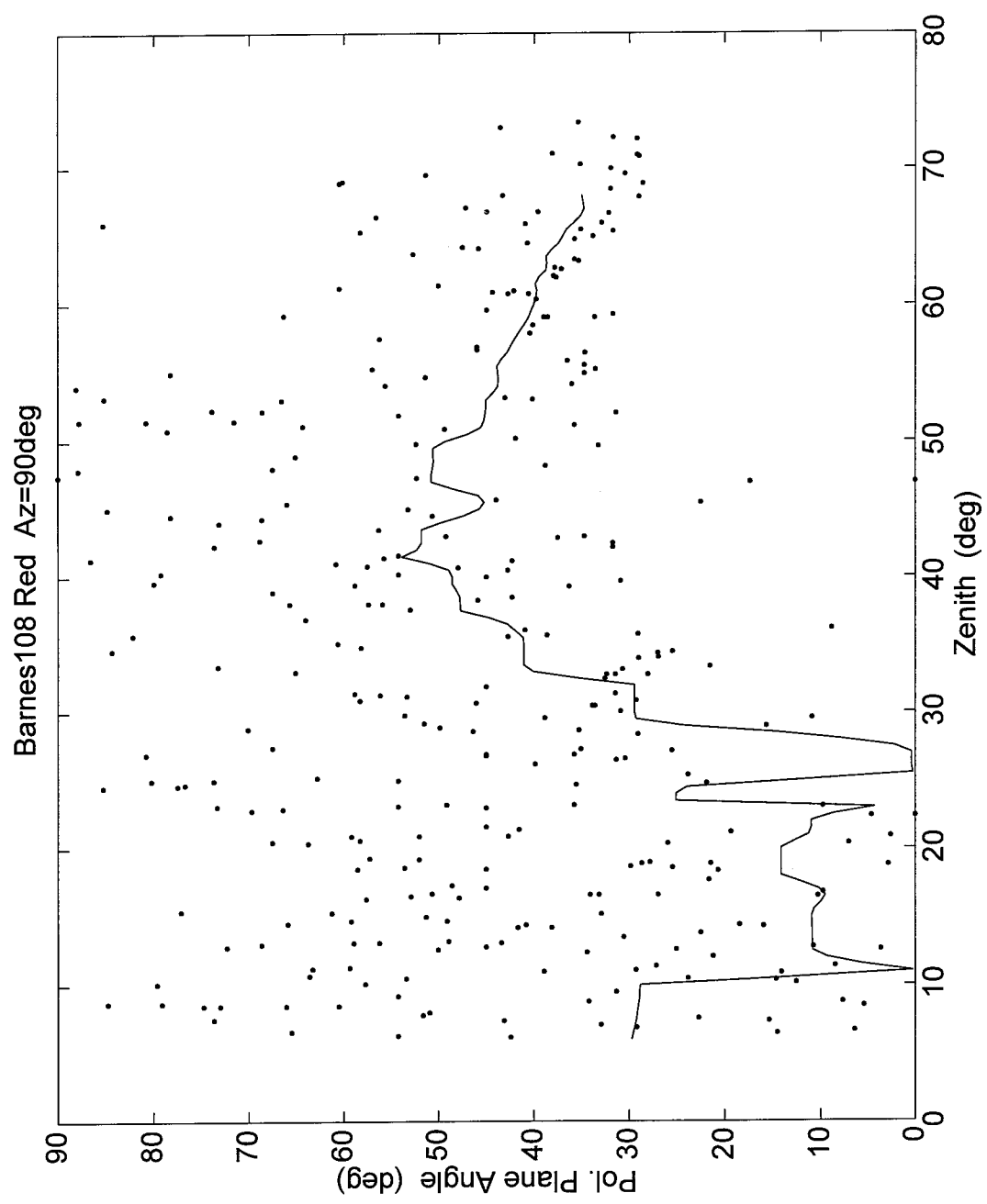


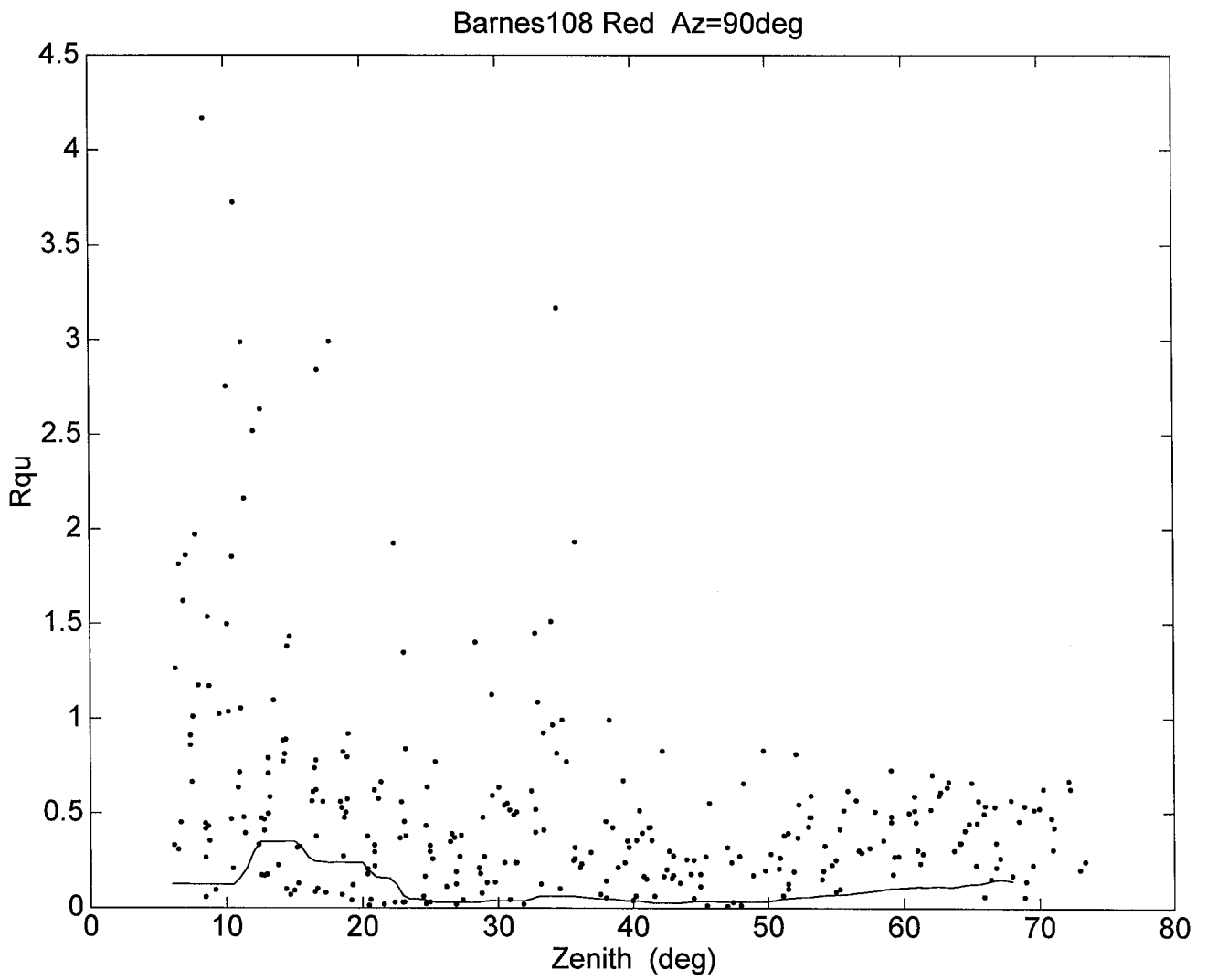


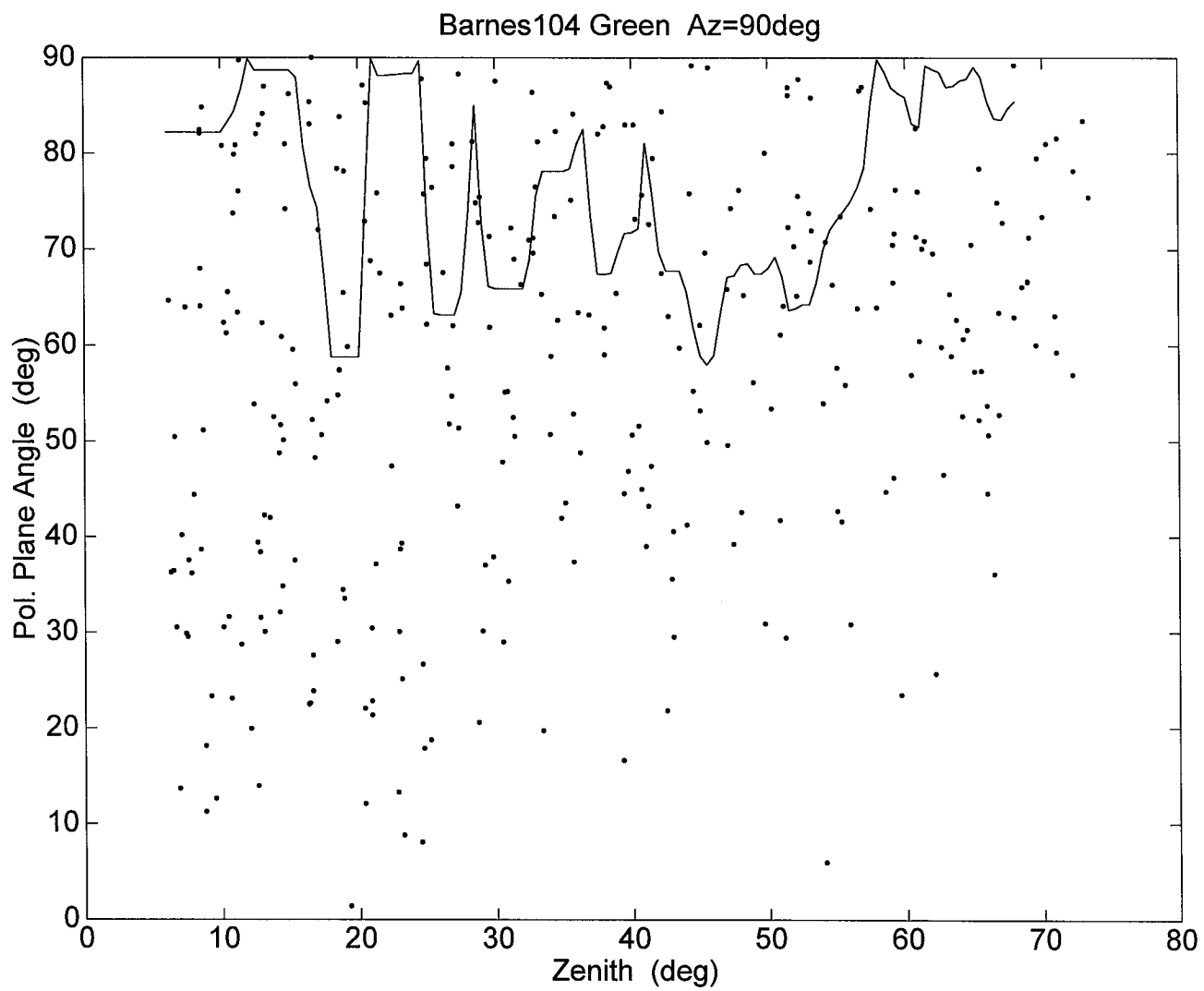


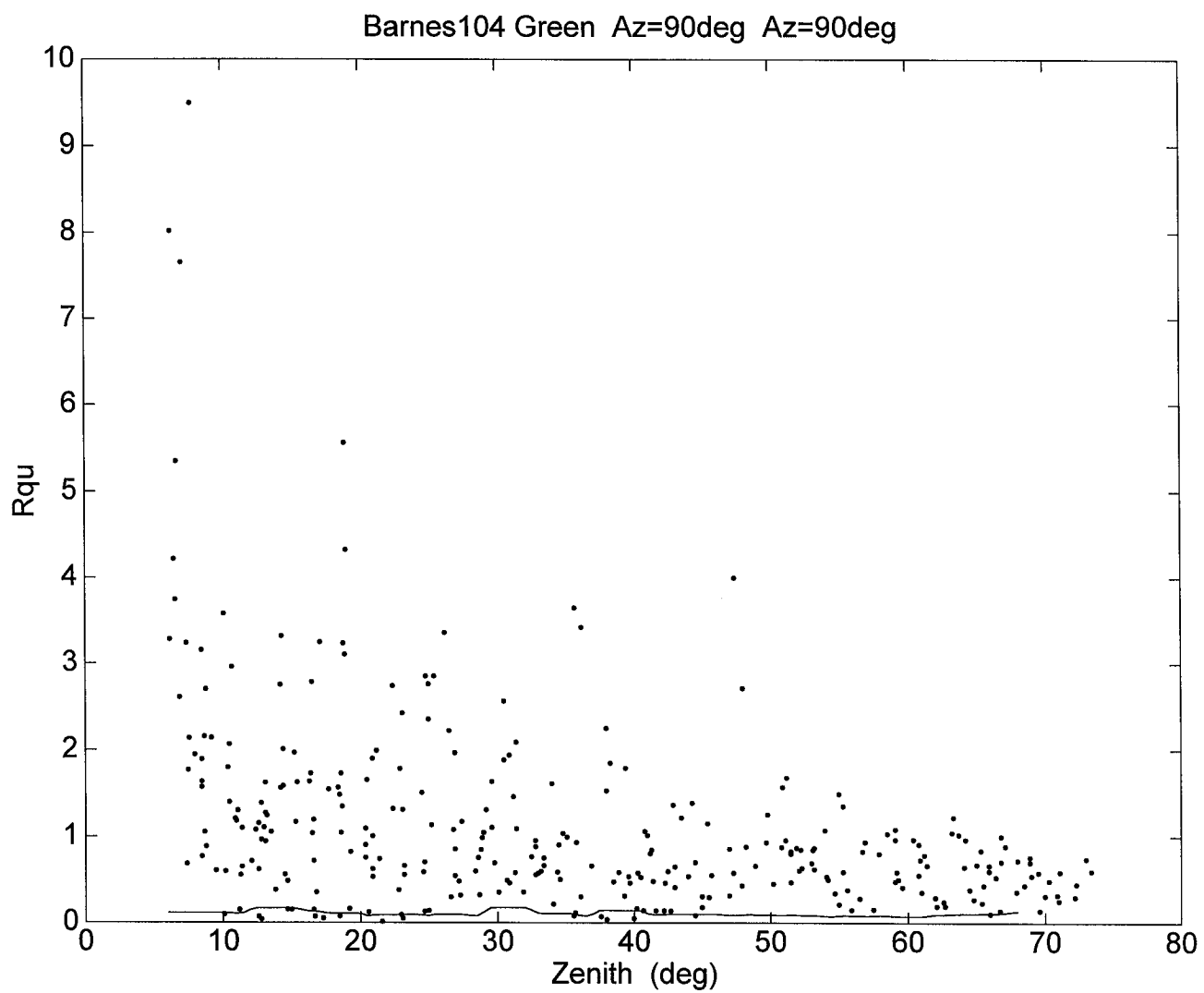




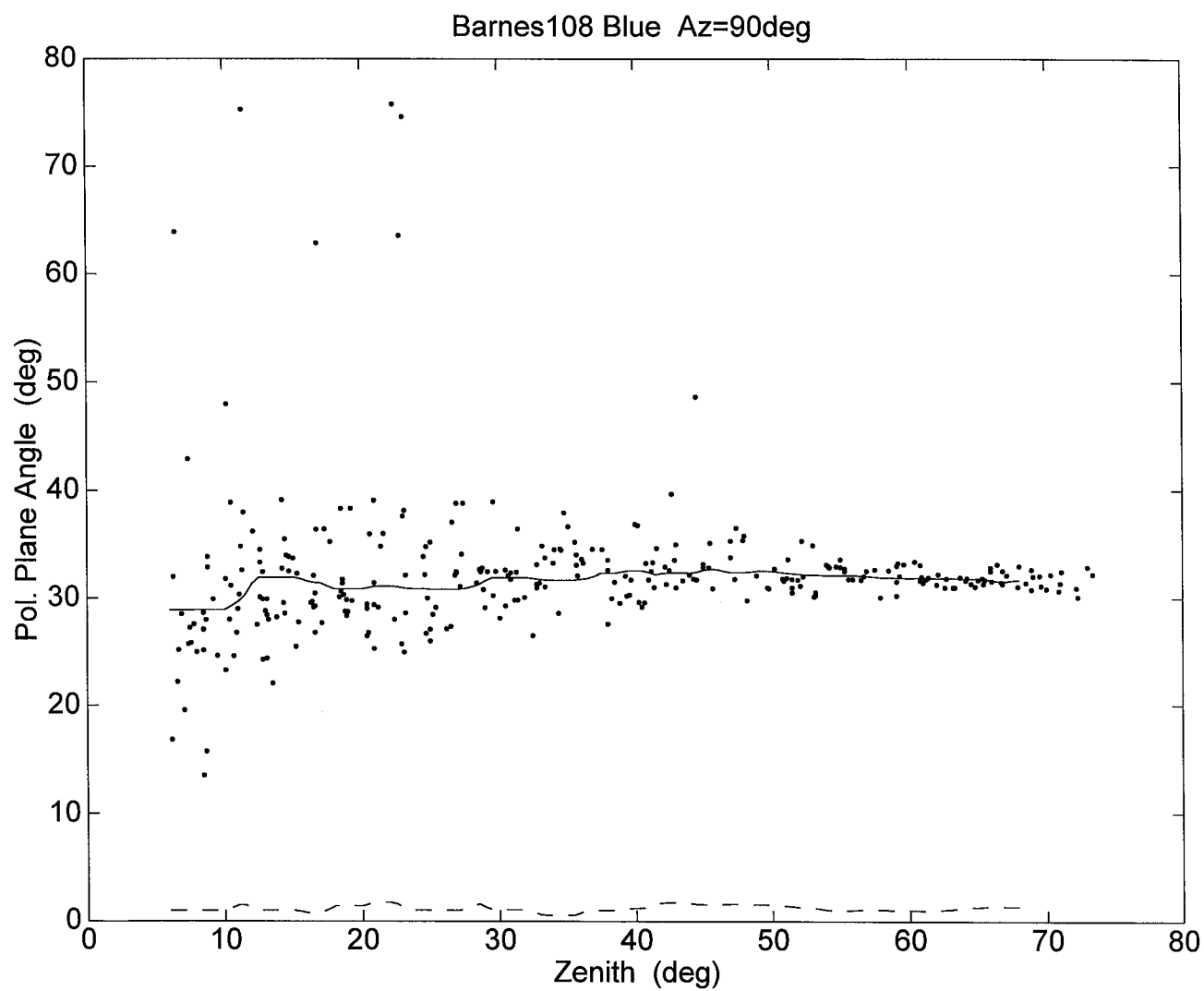








45



47

